
Article:
Johnson, J orcid.org/0000-0003-0431-013X, Panagioti, M, Bass, J et al. (2 more authors) (2017) Resilience to emotional distress in response to failure, error or mistakes: A systematic review. Clinical Psychology Review, 52. pp. 19-42. ISSN 0272-7358

https://doi.org/10.1016/j.cpr.2016.11.007

RUNNING HEAD: Resilience to emotional distress in response to failure

Resilience to emotional distress in response to failure, error or mistakes: A systematic review

Published in Clinical Psychology Review

Accepted November 2016

Judith Johnson*[ab], Maria Panagioti[c], Jennifer Bass[a1], Lauren Ramsey[a2], and Reema Harrison[d3]

*Corresponding author; j.johnson@leeds.ac.uk; Tel: +44 (0)113 3435724; Fax: +44 (0)113 3435749

a. School of Psychology, University of Leeds, Leeds, LS29JT, UK

b. Bradford Institute of Health Research, Bradford Royal Infirmary, Bradford, BD9 6RJ, UK

c. NIHR School for Primary Care Research and Manchester Academic Health Science Centre, University of Manchester, Manchester, M13 9PL, UK; maria.panagioti@manchester.ac.uk

d. School of Public Health, University of Sydney, Sydney 2006, Australia

1. Present address: Cardiology Department, E floor, Jubilee Wing, Leeds General Infirmary, Leeds, LS1 3EX; jenniferbass@nhs.net

2. Present address: Inhealthcare, 58 Netherwood Avenue, Castleford, WF10 2QW, UK; lauren.ramsey@inhealthcare.co.uk

3. Present address: School of Public Health & Community Medicine, Samuels Building, University of New South Wales, Sydney 2052, Australia; reema.harrison@unsw.edu.au

Resilience to emotional distress in response to failure, error or mistakes: A systematic review

Judith Johnson, Maria Panagioti, Jennifer Bass, Lauren Ramsey, Reema Harrison

Abstract

Perceptions of failure have been implicated in a range of psychological disorders, and even a single experience of failure can heighten anxiety and depression. However, not all individuals experience significant emotional distress following failure, indicating the presence of resilience. The current systematic review synthesised studies investigating resilience factors to emotional distress resulting from the experience of failure. For the definition of resilience we used the Bi-Dimensional Framework for resilience research (BDF) which suggests that resilience factors are those which buffer the impact of risk factors, and outlines criteria a variable should meet in order to be considered as conferring resilience. Studies were identified through electronic searches of PsycINFO, MEDLINE, EMBASE and Web of Knowledge. Forty-six relevant studies reported in 38 papers met the inclusion criteria. These provided evidence of the presence of factors which confer resilience to emotional distress in response to failure. The strongest support was found for the factors of higher self-esteem, more positive attributional style, and lower socially-prescribed perfectionism. Weaker evidence was found for the factors of lower trait reappraisal, lower self-oriented perfectionism and higher emotional intelligence. The majority of studies used experimental or longitudinal designs. These results identify specific factors which should be targeted by resilience-building interventions.

Resilience; failure; stress; self-esteem; attributional style; perfectionism

Introduction

Impact of failure experiences

A large body of research suggests that experiencing failure has marked emotional and psychological consequences across a range of individuals and settings. Longitudinal studies indicate that academic failure in adolescents increases risk for clinical depression in adulthood (McCarty et al., 2008; Reinherz, Giaconia, Hauf, Wasserman, & Silverman, 1999), and in those who are depressed, perceived failure has been associated with suicide attempts (Bulik, Carpenter, Kupfer, & Frank, 1990). Even a single experience of failure in non-clinical groups can have significant emotional sequelae. In athletes, match failure has been linked with elevated feelings of depression, humiliation and guilt (Jones & Sheffield, 2007; Wilson & Kerr, 1999), and in healthcare professionals, involvement in medical errors or patient safety failures is reported to result in feelings of shame, depression and anxiety, which can then increase the risk of further errors (Sirriyeh, Lawton, Gardner, & Armitage, 2010; West, Tan, Habermann, Sloan, & Shanafelt, 2009). The reliable impact of failure experiences on mood makes false failure feedback tasks suitable for use as negative mood inductions in experimental settings (Nummenmaa & Niemi, 2004). Studies employing these tasks have found that manipulated failure feedback consistently increases feelings of sadness, defeat and frustration (Johnson, Gooding, Wood, Taylor, & Tarrier, 2011a; Johnson, Tarrier, & Gooding, 2008b; Nummenmaa & Niemi, 2004) and may have a detrimental impact upon cognitive functioning such as reducing the accuracy of memory recall (Johnson et al., 2008b).

However, not all individuals experience significant emotional distress in response to failure, and several psychological models highlight the role of psychological responses to failure in the development of failure-related distress and emotional disorder. For example, cognitive models of suicide have emphasised the role of situation appraisals, suggesting that suicidal thoughts occur when individuals appraise their circumstances in terms of failure (termed ‘defeat’) and entrapment (Johnson, Gooding, & Tarrier, 2008a; Williams, 1997). Yet such models have been criticised for their acceptance of an overly negative, disorder-based

approach to understanding mental health (Johnson & Wood, 2016). By focusing on the development of mental health problems rather than mental wellbeing, it has been suggested that such approaches fail to identify and capitalise on natural coping mechanisms (Johnson & Wood, 2016). As such, they may be missing potential points for psychological interventions to target and develop.

Resilience-based approaches

An alternative to these models are resilience-based approaches (Bonanno, 2004; Masten, 2001; Masten & Powell, 2003). These aim to understand the factors that enable individuals to withstand stressors and avoid psychological distress rather than focusing on the mechanisms that lead to distress and disorder. Resilience-based approaches have the potential to highlight skills and tendencies that individuals can develop to maintain psychological health, leading to a more positively oriented approach to wellbeing. However, this body of literature has suffered from two main limitations.

First, there has been a lack of clarity concerning the criteria for identifying a 'resilient' outcome. The common definition of resilience as factors which *reduce negative outcomes in the face of adversity* would suggest that resilience variables are those which moderate or attenuate the association between risk factors and negative outcomes. In contrast, many studies of resilience have used a correlational approach. These studies have assumed that resilience variables are those which are 'positive', and have investigated whether high levels of a proposed resilience variable (e.g., high perceived social support) is directly associated with lower levels of a negative outcome (e.g., suicidal thoughts). However, as has been highlighted elsewhere (Johnson & Wood, 2016; Johnson, Wood, Gooding, Taylor, & Tarrier, 2011b), every negative variable exists on a continuum with its positive inverse. Returning to the above example, using this approach, it could just as easily be suggested that low perceived social support is a risk factor for suicidal thoughts.

Second, this research failed to lead the field towards more nuanced understandings of resilience. A common approach has been to propose a concept of resilience, develop a questionnaire to measure this, and to investigate the association of this variable in relation to

various outcome variables in different populations. This approach does not enable the proposed resilience variable itself to evolve in order to accommodate new research findings. Indeed, despite fifty years of resilience research, key questions regarding the nature of resilience remain, which may be linked to the limitations of this approach. These concern i) whether factors which confer resilience vary depending on the outcome under consideration (i.e., whether resilience to general mental wellbeing is similar to resilience to negative behavioural outcomes such as suicidality), and ii) whether factors which confer resilience vary according to the risk factor/adversity individuals are facing.

In line with these broader limitations, despite a large growth of interest in resilience, and an increasing awareness of the emotional impact of failure experiences, very few studies have aimed to investigate resilience to emotional distress in response to failure in particular. Of the two studies we identified which have explicitly focused on this, the first investigated whether learning orientation buffered state self-esteem in response to a test result in students, but no significant effect was found (Niiya, Crocker, & Bartmess, 2004). The second investigated the impact of explanatory style on response to sports failure in children, using heart rate acceleration as an indicator of emotional arousal (Martin-Krumm, Sarrazin, Peterson, & Famose, 2003). This suggested that individuals with a pessimistic explanatory style showed a greater increase in heart rate following failure than individuals with an optimistic explanatory style. However, all individuals in the study were exposed to failure, and no analyses investigated whether explanatory style acted as a buffer or moderator of the association between failure and heart rate response.

Given that failure and failure-related distress have been implicated in the development of a range of mental health disorders (Bulik et al., 1990; Johnson et al., 2008a; Reinherz et al., 1999), a fuller and more detailed understanding of resilience in relation to failure could have important implications for psychological interventions. This knowledge could be particularly important for groups likely to experience significant failure events in their occupations, such as health professionals, most of whom will be involved in patient safety failure and clinical errors during their career (Sirriyeh et al., 2010). It could also be

pertinent for young adults in the education system, which has been criticised for becoming increasingly assessment focused (Putwain, 2008), with the pressure of failure cited as contributing to increasing rates of mental health problems in this group (McManus, Bebbington, Jenkins, & Brugha, 2016).

The Bi-Dimensional Framework for Resilience Research

The Bi-dimensional Framework for investigating resilience (BDF; Johnson, 2016; Johnson et al., 2014; Johnson et al., 2011b) was proposed to address these criticisms of the field of resilience research, and to enable the development of evidence-based concepts of resilience. The BDF outlines clear criteria that a variable should meet in order to be considered as conferring resilience. In line with common definitions of resilience, it suggests that resilience factors are those which interact with (or statistically moderate) the likelihood that risk will lead to negative outcomes (Johnson et al., In press). Individuals who are low on resilience will show increasing evidence of negative outcomes with increasing risk, but high resilience individuals will maintain low levels of a given negative outcome, despite risk exposure (see Figure 1). As such, it purports that any investigation of resilience should include three variables, i) the risk factor, ii) the resilience factor, and iii) the outcome variable, and studies directly investigating associations between a predictor variable and an outcome are insufficient to establish a resilience effect. In line with the observations that all variables lie on a continuum from positive to negative, the BDF proposes that all factors can be viewed as ‘bipolar’, and whether they are framed in positive or negative terms is essentially arbitrary (see Figure 2). As such, unlike previous resilience approaches the emphasis of the BDF is not upon identifying ‘positive’ factors which are inversely linked with negative outcomes, but upon identifying psychological factors which can alter the impact of risk.

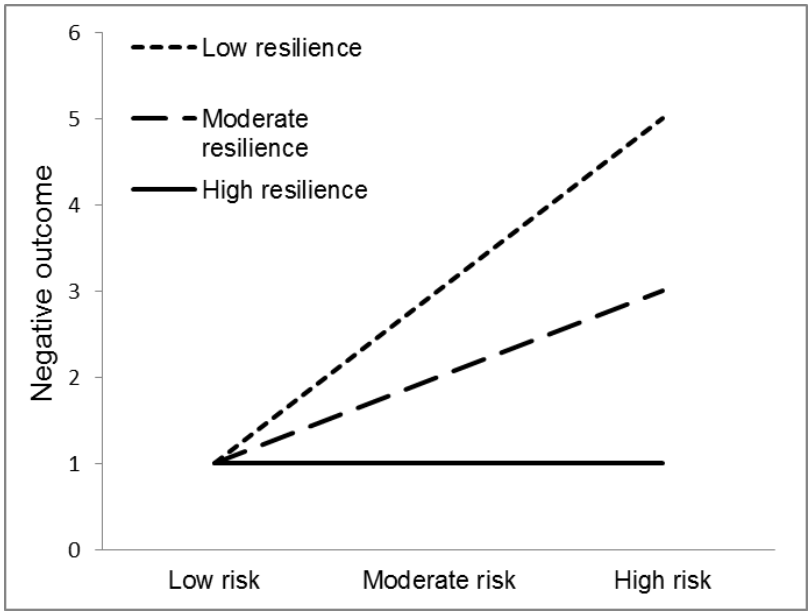

Figure 1.

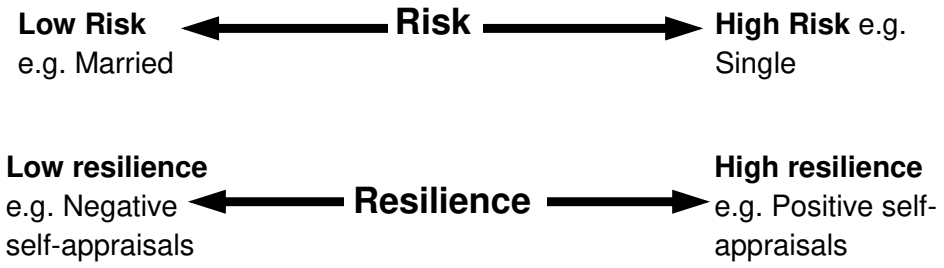

Figure2.

A particular strength of the framework is that it offers a way to aggregate and review existing studies based on i) a particular outcome of interest (e.g., emotional/behavioural outcome), ii) whether a psychological factor has been included, and iii) whether a psychological factor has been examined as a moderator of a risk factor. Importantly, studies that meet these criteria may not have self-identified as having investigated 'resilience'. As such, although there have been few studies which have explicitly aimed to investigate resilience to failure, by using the framework, it is possible to define failure experiences as the risk variable of interest, measures of emotional distress as the outcome variable, and psychological factors as the potential resilience variable, and to use these terms to search

the literature. This approach offers a systematic route to identifying factors which confer resilience to emotional distress/dysfunction in response failure. Given the centrality of emotional distress most mental health disorders, results from this review could have broad relevance to psychological interventions. The BDF was initially developed to investigate resilience to suicidality (Johnson, Gooding, Wood, & Tarrier, 2010a; Johnson et al., 2010b), and underpinned a systematic review in this area. This review suggested that attributional style, sense of agency and lower perfectionism and hopelessness conferred resilience. However, risk factors investigated in these studies varied and only two investigated resilience to failure, neither of which reported on emotional distress/dysfunction as an outcome, instead focusing on suicidal related thoughts (Priester & Clum, 1992, 1993).

Objectives

We undertook a systematic review and evidence synthesis on resilience to failure and error, aiming:

- To investigate whether there are psychological factors which confer resilience to emotional distress in response to failure, error and mistakes
- To evaluate and compare the evidence for different types of psychological variables in conferring resilience

Methods

Methods and results are reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009).

Protocol and registration

The review was registered with the PROSPERO International prospective register of systematic reviews, DOI: 10.15124/CRD42015026761. It is available online at http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015026761.

Search strategy

Four electronic bibliographic databases were searched (from inception to September 2014, and then updated to April 2016): PsycInfo, Ovid Medline, EmBase and Web of Knowledge. We searched for papers containing at least one term from each of the following blocks: (fail* or error* or defeat or mistake*) and (interact* or moderat* or buffer* or amplif*) and (anxiety or anxious or depression or depressed or emotion* or affect or mood or shame or guilt or PTSD or trauma or insomnia). A combination of Medical Subject Headings (MESH terms) and text words were used in our searches (see Supplementary File 1 for the Medline search strategy). No previous reviews were identified in the area.

Eligibility criteria

Studies were eligible for inclusion if they met the following criteria:

- Population: We included studies which were conducted among adults.
- Setting: Our focus was not restricted to studies conducted in a particular setting, such as healthcare or educational settings.
- Design
    - Quantitative research designs. We included studies with any type of quantitative research design ranging from experimental studies to observational studies (cohort and cross-sectional studies).
    - We included studies which examined moderators of the association between error/failure and emotional distress/dysfunction, or factors which interact with the experience of error/failure to predict psychological outcome (using moderated regression or other statistical methods of investigating two-way interactions).
- Outcome measure
    - We included studies which reported data on outcome measures of emotional distress or dysfunction which could encompass a range of outcomes such as general positive and negative affect, depression/depressive symptoms, anxiety and self-esteem (Ridner, 2004).

- Resilience variable
    - Variables in the moderation/interaction analysis could be regarded as a potential 'psychological resilience factor', i.e., a psychological quality of individuals, such as a belief, tendency or ability.
    - As the review was interested in naturally occurring resilience, studies where resilience variables had been manipulated via experimental manipulation were excluded.
- Experience of error or failure
    - We included studies where some or all of the participants experienced error or failure, either naturally occurring or experimentally manipulated.

Exclusion criteria:

- Studies that were not in the English Language, did not involve human participants and grey literature studies were excluded.
- Studies which only investigated 3-way interactions were excluded, as the relationships tested in these studies were very complex.
- As demographic factors and clinical disorders (including narcissism) are not considered potential resilience factors by resilience frameworks, studies of these variables were excluded.
- Due to the complex nature of social interactions, and the range of causes that can contribute to relationship breakdown, studies of social rejection or perceived social failure were excluded.
- Studies where participants only imagined failure events were excluded.

Study selection

Initially 20% of the titles/abstracts were screened by three reviewers independently to reach consensus within the team regarding the study selection criteria (JJ;JB; LR). All the remaining titles/abstracts were screened independently by two of these reviewers. The full

texts of studies assessed as potentially eligible for the review were then retrieved and checked against the inclusion and exclusion criteria by two researchers working independently (JJ and MP or RM). Any disagreements were resolved by discussion.

Data extraction

A data extraction table was devised in Microsoft Excel and initially piloted on five studies. We extracted the following descriptive data: country, year of publication, participant characteristics (population, number, mean age, percentage male), research design, statistical analysis conducted, proposed resilience variable, failure/error type or manipulation, outcome variable, key results of the interactions (moderation analyses), and critical appraisal information. Data were extracted by the first author, with any uncertainties addressed in discussion with the second author.

Risk of bias assessment

The majority of the studies included in the review were experimental studies, with observational cross-sectional and longitudinal studies also included. As well as distinguishing between these different designs, we also assessed for the following risk of bias criteria:

1. Whether measures of the resilience and outcome variables used validated questionnaires
2. Whether the statistical analysis controlled for confounders (e.g., baseline levels of the outcome measure/s)
3. Whether response rate or data capture among eligible participants was recorded and found to be at 70% or greater at baseline
4. Whether response rate was recorded and found to be at 70% or greater at follow-up (for longitudinal studies only)
5. Whether participants were randomly assigned to conditions (for experimental studies only)

6. Whether random assignment was based on random sequence generation (for experimental studies only)
7. Whether use of allocation concealment to conditions was employed (for experimental studies only).

These criteria were based upon Cochrane risk of bias criteria (Higgins & Green, 2008) and guidance for the assessment of observational studies (CRD, 2009) . Studies were assigned a rating of 1 for each criterion met (maximum rating of 4 for cross-sectional observational studies, 5 for longitudinal studies and 7 for experimental studies).

Data synthesis

Assessment of the strength of the moderating impact of potential resilience variables between failure and emotional distress through meta-regression would have been desirable (Schmidt & Hunter, 2014). However, this was not possible due to wide heterogeneity between studies regarding the measurement of the emotional distress outcome. A narrative synthesis was therefore undertaken, which integrated review results in a non-quantitative but connected way (Keeley, Storch, Merlo, & Geffken, 2008; Knopp, Knowles, Bee, Lovell, & Bower, 2013). Where more than one study had investigated the same proposed resilience variable, we used a box-score approach. In the box-score approach, the relationship between moderating variables and outcomes is tabulated in terms of significance and direction (negative, positive, or no relationship) (Green & Hall, 1984). Studies within each respective group were tallied and the majority of studies within any specific category was considered to indicate the likely relationship between the potential resilience variable and the outcome (Light & Smith, 1971). The advantages of the box-score approach were that it enabled basic quantification of reported moderator effects and identification of patterns across collated studies. It also enabled quantification of the relationship between quality of analyses and reported effects.

Results

Overall, 5071 titles and abstracts were screened for eligibility. Following screening, 38 papers (reporting 46 relevant studies) met the inclusion criteria (see PRISMA flow chart displayed in Figure 3).

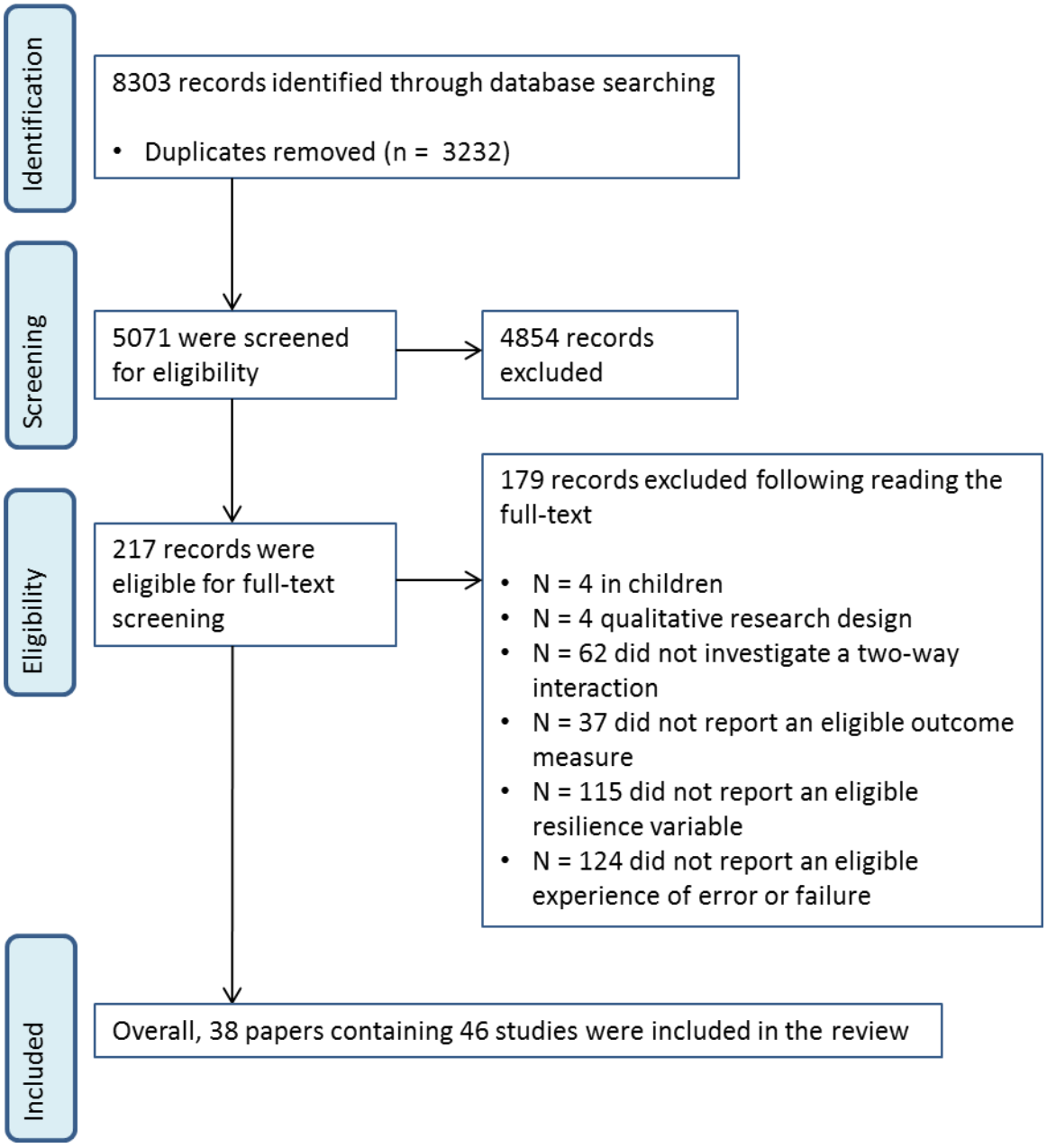

Figure 3.

Characteristics of studies and populations

Included studies had a total of 5905 participants (*m* sample size = 128.37, *SD* = 83.8, range = 46-399). The mean age of participants ranged from 18.6 to 47.6 (data missing for 25 studies), and the majority of studies were amongst undergraduate students, with only one

study conducted in a clinical population (Johnson et al., 2011a). The gender split varied across studies, but overall participant groups comprised slightly more females (*m* = 38.6% male participants, *SD* = 22.6, data missing for 5 studies). Most studies were experimental (80.43%), with the remainder using longitudinal (15.22%) and cross-sectional (4.35%) designs. Studies were from a range of countries, but a large proportion (60.87%) was conducted in the USA.

Tables 1 and 2 here

Characteristics of resilience, failure and emotional distress variables

A number of potential resilience factors were investigated. The single factor most frequently studied was self-esteem (see Table 3 for the box-score review), with a total of 15 studies (32.6%) investigating this. Other factors investigated a range of trait coping and personality constructs, such as attributional style, emotional intelligence, perfectionism and reappraisal. Resilience factors were measured using validated questionnaires in 40 studies (87%). Similarly, a range of emotion distress outcome variables were studied, including depression (*n*=13, 28.3%), anxiety (*n*=8, 17.4%), general affect (*n*=18, 39.1%) and negative self-relevant emotions (*n*=9, 19.6%). Thirty studies (62.5%) used validated questionnaires to measure the outcome variable (or at least one outcome variable, where more than one was investigated), 18 (60%) of which reported significant results. Of these, nine studies used validated measures of depressed or anxious mood, such as the Beck Depression Inventory (Beck, 1967), the State–Trait Anxiety Inventory (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) and the Multiple Affect Adjective Checklist (Zuckerman & Lubin, 1965), and 6 (66%) reported significant results.

In order to study reactions to failure, the majority of studies used an experimental approach and a false failure paradigm. In these paradigms the task is fixed to be too difficult to pass, the feedback received by participants is fixed to report failure regardless of performance, or a combination of both of these are applied. The single most common false failure task used was the Remote Associates Task (RAT) (Mednick, 1962) or an adapted version of this (*n* = 12, 26.1%). In the RAT task, participants guess a target word from three

indicator words which are fixed to be easy or difficult in order to lead to failure or success. Purported intelligence tests were used to induce failure in six (13%) studies and insoluble anagram tasks were used in five (10.9%) studies. Of the eight longitudinal studies, seven investigated reactions to exams or academic grades, and one investigated acceptance or rejection to university.

Table 3 here

Risk of bias assessment

The results of the risk of bias assessment are displayed in Figure 4. Out of a total possible score of 7, experimental studies scored between 0 and 4 (m = 2.95, SD = .91). Whilst 89.2% of these studies used a validated questionnaire for the resilience variable and 81.1% used random assignment (with this variable not applying to an additional 5.4% of studies which used repeated measures), fewer (54.1%) controlled for confounders such as baseline mood. Furthermore, no studies reported whether they used random sequence generation, and only 1 reported using allocation concealment. Out of a total possible score of 5, longitudinal studies scored between 2 and 4 (m=2.86, SD=.69). All studies used a validated resilience questionnaire, and most (71.4%) controlled for confounders and used a validated emotional distress outcome questionnaire. However, few (28.6%) reported the response rate at follow-up and found this to be $\geq$70%. There were two cross-sectional studies with a maximum possible score of 4. One of these studies, one scored 2, the other scored 0.

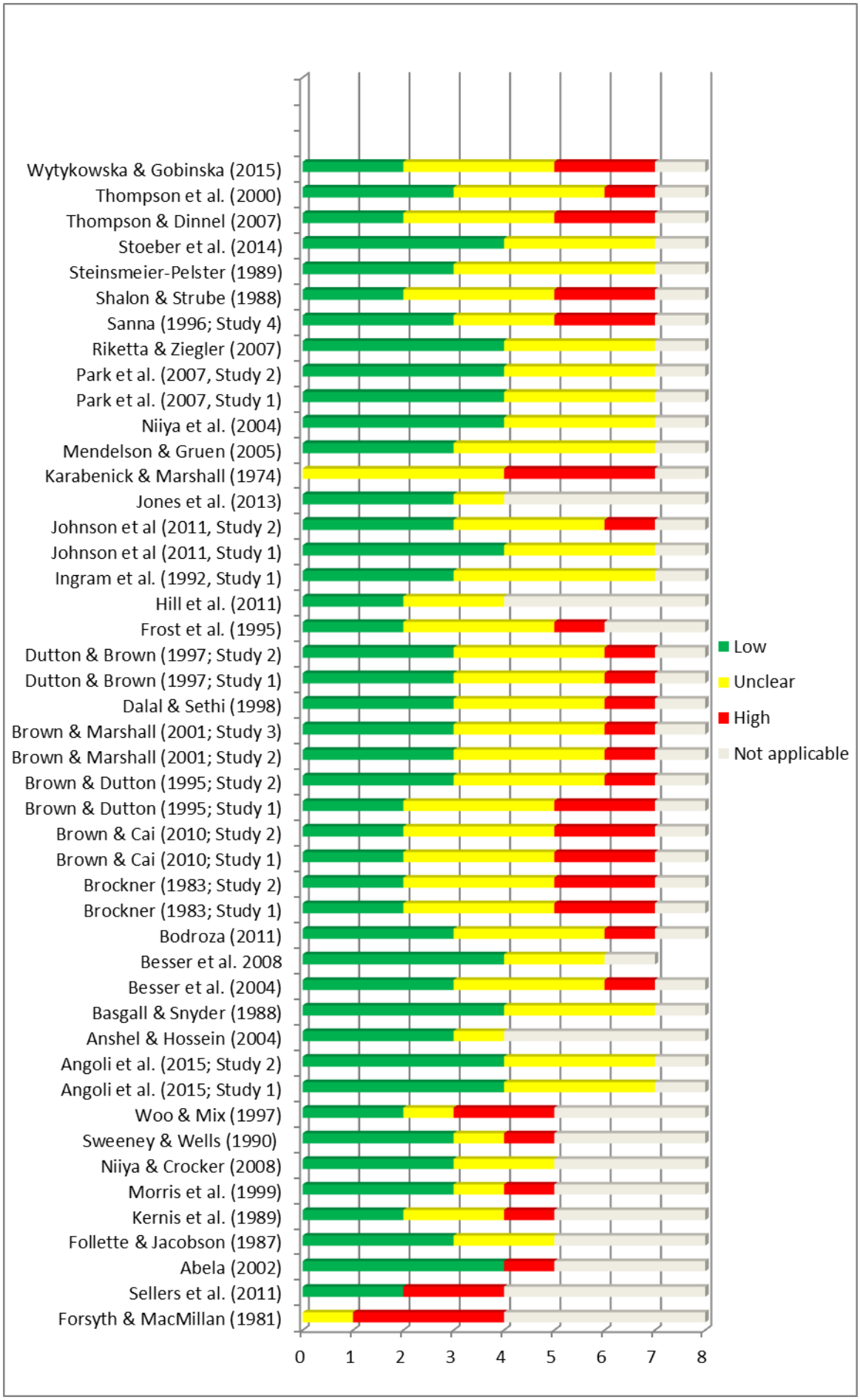
Wytykowska & Gobinska (2015)
Thompson et al. (2000)
Thompson & Dinnel (2007)
Stoeber et al. (2014)
Steinsmeier-Pelster (1989)
Shalon & Strube (1988)
Sanna (1996; Study 4)
Riketta & Ziegler (2007)
Park et al. (2007, Study 2)
Park et al. (2007, Study 1)
Niiya et al. (2004)
Mendelson & Gruen (2005)
Karabenick & Marshall (1974)
Jones et al. (2013)
Johnson et al (2011, Study 2)
Johnson et al (2011, Study 1)
Ingram et al. (1992, Study 1)
Hill et al. (2011)
Frost et al. (1995)
Dutton & Brown (1997; Study 2)
Dutton & Brown (1997; Study 1)
Dalal & Sethi (1998)
Brown & Marshall (2001; Study 3)
Brown & Marshall (2001; Study 2)
Brown & Dutton (1995; Study 2)
Brown & Dutton (1995; Study 1)
Brown & Cai (2010; Study 2)
Brown & Cai (2010; Study 1)
Brockner (1983; Study 2)
Brockner (1983; Study 1)
Bodroza (2011)
Besser et al. 2008
Besser et al. (2004)
Basgall & Snyder (1988)
Anshel & Hossein (2004)
Angoli et al. (2015; Study 2)
Angoli et al. (2015; Study 1)
Woo & Mix (1997)
Sweeney & Wells (1990)
Niiya & Crocker (2008)
Morris et al. (1999)
Kernis et al. (1989)
Follette & Jacobson (1987)
Abela (2002)
Sellers et al. (2011)
Forsyth & MacMillan (1981)
0
1
2
3
4
5
6
7
8
Low
Unclear
High
Not applicable

Figure 4.

Are there factors which confer psychological resilience to emotional distress in response to failure?

The review identified a number of studies which reported psychological variables which interact with experiences of failure, errors or mistakes in order to predict mood. Notably, there were eight potential resilience variables which were tested in more than one study (see Table 3 for a box-score review of these). Four of these (self-esteem, attributional style, socially prescribed perfectionism and trait reappraisal) were found to significantly moderate the association between failure and emotional distress in >50% of the studies in which they were tested, two drew a balance of significant and null findings (self-oriented perfectionism and emotional intelligence) and two drew only null findings (academic self-worth and trait emotion suppression).

Which potential resilience factors have the most supporting evidence?

Of the four potential resilience variables with the most supporting evidence, three drew significant results in two-thirds of the studies which tested them (self-esteem, attributional style and socially prescribed perfectionism). Self-esteem was the most frequently tested of these. It was investigated in three longitudinal studies (all of which reported a significant moderation effect) and 12 experimental studies (seven of which reported a significant moderation effect). Twelve studies measured self-esteem using the same measure, the Rosenberg Self-Esteem Scale (Rosenberg, 1965). Two of these included a validated measure of depressed mood as the outcome (Abela, 2002; Sweeney & Wells, 1990). Both of these were longitudinal studies of reactions to naturally occurring failure. The remaining self-esteem studies used a range of mood and affect measures, including the Feelings of Self Worth Scale (Brown & Dutton, 1995) which measures the extent to which participants are proud, pleased with themselves, humiliated and ashamed. The risk of bias score of studies which reported significant results (m=2.5) was similar to that of studies reporting non-significant results (m=2) suggesting that quality variation is unlikely to have affected significance of findings.

Attributional style was tested in six studies, including three experimental studies (two of which reported significant results), two longitudinal studies (one of which reported a significant moderation effect) and one cross-sectional study (which reported a significant moderation effect). Three studies (Follette & Jacobson, 1987; Morris & Tiggemann, 1999; Stiensmeier‐Pelster, 1989) used a version of the Attributional Style Questionnaire (ASQ; Peterson et al., 1982), two used a single item (Brown & Cai, 2010), and one used an non-validated three-item scale (Forsyth & McMillan, 1981). The three which did not use a version of the ASQ asked about attributions for a specific event. Validated questionnaires of depressed mood were used to measure the emotional distress outcome in two studies, with the remainder using the Feelings of Self Worth Scale (Brown & Dutton, 1995), Visual Analogue Scales (one study) and a measure of general affect (one study). The risk of bias score of studies reporting significant results (m=1.75) was similar to that of studies reporting non-significant results (m=2).

Socially-prescribed perfectionism was tested in three studies which each used an experimental design; two of these reported a significant interaction. Two studies measured the emotion outcome with scales developed for the study (one of which reported a significant interaction), and the third measured the emotion distress outcome with validated measures of anxiety, depression and anger. This third study found significant interactions for each of these emotion outcomes. The pattern of interactions was such that lower levels of perfectionism were protective against emotional distress in response to failure. The risk of bias scores of the three studies were similar (the study reporting non-significant results scored three, compared to a score of four for the two remaining studies).

Only one potential resilience variable, trait reappraisal was found to interact with failure in each study in which it was tested (Johnson et al., 2011a), but this is may be due to the small number of studies in which this was included (two in total). Emotional distress outcomes were measured using visual analogue scales (both studies) and a validated measure of general affect (one study). The pattern of the interactions was such that lower

levels of trait reappraisal buffered individuals from higher levels of negative mood in response to failure. Conversely, two variables (self-oriented perfectionism and emotional intelligence) drew equivocal findings and two (academic self-worth and trait emotion suppression) were not significant moderators of failure in any of the studies in which they were tested.

Three-way interactions between two resilience variables and failure

In four studies reported in three papers (Abela, 2002; Niiya & Crocker, 2008; Park, Crocker, & Kiefer, 2007), results from the two-way interactions between potential resilience variables and failure were qualified by significant three-way interactions involving a second potential resilience variable (see Supplementary File 2). These interactions suggested that the moderating impact of one proposed resilience variable on emotional response to failure varies depending on the degree of another proposed resilience variable. In three of the four studies, self-esteem was included as one of the resilience variables. Together, these results suggest that the moderating impact of self-esteem on emotional response to failure varies according to pessimism and the extent to which self-worth is contingent on academic performance. In particular, individuals with either pessimistic inferential style or higher contingencies of self-worth in combination with low-self-esteem were more vulnerable to low mood in response to failure.

Discussion

The first objective of the current review was to investigate whether there are psychological constructs which can buffer the association between experiences of failure, errors or mistakes, and emotional distress or dysfunction. The second objective was to identify specific psychological factors which may have this buffering effect, and which can be regarded as conferring resilience to failure. The review used the Bi-dimensional Framework for resilience research (BDF; Johnson et al., 2011b) which proposes that resilience factors are those which statistically moderate the likelihood that risk factors, such as failure experiences, will lead to negative outcomes such as emotional distress.

Summary of findings

The review found clear evidence for the existence of psychological factors which buffer the association between failure experiences and emotional distress or dysfunction. A range of personality and coping constructs were investigated, and the strongest support was found for the factors of higher self-esteem, more positive attributional style and lower levels socially prescribed perfectionism. Several other variables had a weaker evidence base due to smaller number of studies or more equivocal results, but may also buffer emotional distress in response to failure. These included lower levels of trait reappraisal, lower self-oriented perfectionism and higher emotional intelligence. Two variables, academic self-worth and trait emotion suppression, were investigated in more than one study but were not found to be significant moderators, suggesting that these do not confer resilience to failure.

Implications for psychological resilience-building interventions for clinical and non-clinical populations

The concept of building resilience has long been an implicit aspect of psychological interventions in populations with psychological disorders. For example, Cognitive-Behaviour Therapy (CBT) aims to help clients develop skills and techniques for managing low mood and stress which they can put into practice in daily life when the need arises (Beck, 1976). Although the focus of the therapy may be on alleviating the client's current distress, an underlying assumption has been that these skills will be a source of resilience for the client after therapy has ceased. Recent years have seen a growing focus on this element of interventions, with therapeutic approaches being developed or refined specifically to prevent subsequent relapses (Williams et al., 2014). There has also been increasing interest in resilience-focused interventions in populations which are not currently experiencing psychological disorder, but may be at heightened risk. These include children and young adults (Dray et al., 2014; Lynch, Geller, & Schmidt, 2004), military families (Saltzman et al., 2011) and healthcare staff (Goldhagen, Kingsolver, Stinnett, & Rosdahl, 2015; Mealer et al., 2014).

These interventions have been designed and developed on the basis of clinical knowledge and factors which predict symptoms over time. However, there has been a lack

of evidence regarding factors which can buffer individuals from emotional distress in response to subsequent stressors, such as failure, which is a strong and consistent trigger of emotional distress (Bulik et al., 1990; Johnson et al., 2011a; McCarty et al., 2008; Reinherz et al., 1999). By identifying factors that these psychological interventions can target in order to reduce risk of emotional distress in response to subsequent failure experiences, results from the review provide an evidence-base for these interventions to draw on. These results are supported by the experimental and longitudinal design of most of the studies, which provide evidence that the proposed resilience variables may have a causational impact on subsequent mood. In particular, the review suggests that resilience-building interventions should aim to increase levels of self-esteem, develop a more positive attributional style, and reduce levels of perfectionism (particularly socially prescribed perfectionism).

In addition to clinical groups, resilience-based interventions could have important implications for groups who may not currently suffering from mental health difficulties, but who are regularly confronted with failure as part of their training or work. One such group are healthcare professionals, who may undertake ongoing training and assessment alongside their practice and who may also be involved in medical errors (Sirriyeh et al., 2010). Research suggests that involvement in medical errors can cause significant emotional distress, and that experiencing distress can then increase the risk of involvement in subsequent errors (Hall, Johnson, Watt, Tsipa, & O'Connor, 2016; Sirriyeh et al., 2010; West et al., 2009). In this group, resilience-based interventions could enable the development of psychological resources which may both reduce emotional distress in response to failure and errors, and improve patient safety.

Comparison with previous findings and Implications for future research

There has been growing interest in the concept of resilience, but the field has suffered from two main limitations which have prevented the development of increasingly advanced and nuanced understandings of resilience. First, there has been a lack of clarity concerning the criteria that a variable should meet in order to be regarded as a resilience factor, and second, the approach to investigating resilience has too often been top-down;

proposing a concept of resilience and then exploring this concept in different settings. This has prevented the natural evolution of concepts of resilience in response to new research findings. Consistent with these limitations, very few studies have sought to investigate resilience to emotional distress in response to failure in particular. Of the two studies we identified which had focused on this topic prior to undertaking the review, neither had reported evidence that a psychological variable conferred resilience to emotional distress in response to failure (Martin-Krumm et al., 2003; Niiya et al., 2004). The current study reviewed the literature using the Bi-Dimensional Framework for resilience research (BDF) which was developed to address limitations in the resilience literature (Johnson, 2016). It suggests that resilience factors are those which statistically moderate or attenuate the association risk factors and negative outcomes, such that at high levels of resilience, the association between exposure to risk factors and negative outcomes is weakened (see Figure 1). This approach identifies relevant studies according to the methodology studies have used, overcoming the terminology used by the authors, and as such allows a broader number of studies to be identified. Using this approach, we found 46 relevant studies, which together drew strong support for the factors of higher self-esteem, more positive attributional style, and lower socially-prescribed perfectionism. Weaker support was drawn for the factors of lower trait reappraisal, lower self-oriented perfectionism and higher emotional intelligence. Given the previous sparsity of research in this area, these results provide a strong foundation for further research into resilience in the face of failure.

These results can also be compared to resilience findings drawn from other areas. Of particular interest is one previous review which used the same framework (the BDF) to synthesise studies investigating resilience to suicidality (Johnson et al., 2011b), identifying attributional style, perfectionism, agency and hopelessness as key buffering factors, with weaker evidence for self-esteem. Factors identified in the current review overlap with these, providing support for these findings and suggesting that factors which confer resilience to suicidality may also buffer individuals from emotional distress in response to failure. The convergence of results is particularly interesting given clear variations between these two

reviews. For example, whereas the previous review included studies investigating a range of risk factors, both internal (e.g., depression) and external (e.g., life stress), with only two studies investigating failure experiences in particular (Priester & Clum, 1992, 1993) the current review focused only on a specific, discrete and external risk factor (failure). Furthermore, whereas the previous review included a number of cross-sectional studies and no experimental studies, the great majority of studies in the current review were of an experimental or longitudinal design. Particularly notable is that no individual study appeared in both reviews. As such, the current review both supports and extends the previous review, providing evidence that self-esteem, attributional style and perfectionism could be key resilience factors for both suicidality and emotional distress which may have a causal role in protecting individuals from the negative impact of failure. In supporting these previous results, the current review also provides further evidence of the utility of the BDF for evidence synthesis. Like the previous review, only a small number of the included studies self-identified as investigations of 'resilience'. However, by using the BDF, methodology was used to select relevant studies instead of terminology, removing this limitation.

The review identified both factors which confer resilience to failure, and those which did not. In particular, academic self-worth and trait emotion suppression were investigated in more than one study but not found to be significant moderators, suggesting that these do not confer resilience to failure. This provides clear indications for factors which future resilience research may build on, and those which can be precluded. Given the similarities between these non-significant variables with those which drew more significant interaction effects (e.g., academic self-worth with self-esteem), conceptual clarity is likely to be important when investigating resilience.

The majority of studies included in the review were experimental, with a smaller number using longitudinal approaches. No studies were identified which investigated resilience using a daily-diary or experience-sampling method. These methods provide a rich data source, allowing for the investigation of associations between resilience factors and day-to-day (or hour-to-hour) fluctuations in mood. Like longitudinal studies, they offer both

evidence regarding causality and an ecologically valid design, but provide a larger number of time points on which to base conclusions. Given that mood can vary dramatically over time, this prevents spurious conclusions being drawn on the basis of one dip in mood, for example. Future resilience research would benefit from extending the current evidence base by using these designs.

The present review took took a systematic approach to investigating resilience to failure, but as it was not a meta-analysis, it was unable to report effect sizes. The main contribution of this review is that it builds the evidence base and supports the formation of specific hypotheses to be tested meta-analytically by future studies. However, in future, a meta-analysis of key supported moderators such as self-esteem is highly encouraged. Such meta-analysis could examine (through meta-regression analysis) the moderating effects of self-esteem on emotional distress independent of whether the primary studies tested interaction effects.

Furthermore, the current study focused on investigating resilience to discrete failure experiences, excluding studies which investigated reactions to perceived social failures and rejections. This decision was made due to the more complex nature of social interactions, which are complex and can have a range of contributors, and which may extend and vary over time. However, given the importance of social relationships to psychological wellbeing and mental health (e.g., Cohen & Hoberman, 1983; Hovey, 1999), investigating resilience to these events may represent an important avenue for future reviews to explore.

Strengths and limitations

The study had several strengths. It is the first systematic review to synthesise literature investigating resilience to failure, and it approached this using a theoretically informed approach. It was conducted and reported according to PRISMA guidelines (Moher et al., 2009). The searches were designed to be comprehensive, and drew a large number of results eligible for inclusion. The majority of included studies used experimental or longitudinal designs which provide some evidence of causality. In all experimental studies apart from two (reported in one paper; Brown & Cai, 2010) the proposed resilience variable

was measured at baseline, preventing the possibility that measurement of these was affected by the failure experience (e.g., Chung et al., 2014).

The review also had limitations. The majority of studies were conducted amongst undergraduate students, and only one study used a clinical population. However, this study was reported in a two-part paper (Johnson et al., 2011a), where the same experiment was repeated in both undergraduate and clinical populations. Results were replicated in both studies, suggesting that the resilience factor (low trait reappraisal) had the same buffering impact in both populations. This provides evidence that although most of these studies were not in clinical populations, results may generalise. Furthermore, nine studies used validated measures of depression and anxiety which are often used in clinical settings (e.g., BDI, State-Trait anxiety inventory) in order to measure the emotional distress outcome. The majority of these found significant results, indicating that the impact of the resilience factors tested by these studies is significant enough to influence clinical levels of mood change.

Study results were aggregated using the box-score approach which allowed for the visual display of significance of findings. A limitation of using this approach was that it was not possible to consider the magnitude of reported effects. As such, it may have led to a more conservative interpretation of the evidence (Green & Hall, 1984; Knopp et al., 2013).

The review only included papers published in peer-reviewed journals. It is now increasingly recognised that grey literature is an additional useful source of research data which can help minimise the possibility of publication and study selection biases in systematic reviews. However, we decided to exclude grey literature from this study because it is very difficult to search, synthesise and appraise the quality of data from grey literature (Mahood, Van Eerd, & Irvin, 2014).

Conclusion

This is the first systematic review to identify resilience factors that may buffer emotional distress or dysfunction resulting from failure, mistakes or errors. Results suggested that higher self-esteem, more positive attributional style and lower levels of socially prescribed perfectionism may confer resilience to emotional distress in response to

failure, and that academic self-worth and trait emotion suppression are not linked with resilience. These results suggest that these factors may be useful targets for resilience-building interventions, and should be incorporated into concepts of resilience. These findings also support the utility of the Bi-Dimensional Framework for the synthesis of studies investigating potential resilience factors.

References

Abela, J. R. (2002). Depressive mood reactions to failure in the achievement domain: A test of the integration of the hopelessness and self-esteem theories of depression. *Cognitive Therapy and Research, 26*, 531-552.

Agnoli, S., Pittarello, A., Hysenbelli, D., & Rubaltelli, E. (2015). "Give, but Give until It Hurts": The Modulatory Role of Trait Emotional Intelligence on the Motivation to Help. *PloS one, 10*, e0130704.

Anshel, M. H., & Mansouri, H. (2005). Influences of perfectionism on motor performance, affect, and causal attributions in response to critical information feedback. *Journal of Sport Behavior, 28*, 99.

Basgall, J. A., & Snyder, C. (1988). Excuses in waiting: External locus of control and reactions to success-failure feedback. *Journal of Personality and Social Psychology, 54*, 656.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Hoeber.

Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. Oxford, England: International Universities Press.

Besser, A., Flett, G. L., & Hewitt, P. L. (2004). Perfectionism, cognition, and affect in response to performance failure vs. success. *Journal of Rational-Emotive and Cognitive-Behavior Therapy, 22*, 297-324.

Besser, A., Flett, G. L., Hewitt, P. L., & Guez, J. (2008). Perfectionism, and cognitions, affect, self-esteem, and physiological reactions in a performance situation. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 26*, 206-228.

Bodroža, B. (2011). Cognitive and affective reactions to success and failure: The quality of feedback as the determinant of self-verifying and self-enhancing motivation. *Psihologija, 44*, 149-166.

Bonanno, G. A. (2004). Loss, Trauma, and Human Resilience: Have We Underestimated the Human Capacity to Thrive After Extremely Aversive Events? *American Psychologist, 59*, 20-28.

Brockner, J., Gardner, M., Bierman, J., Mahan, T., Thomas, B., Weiss, W., Winters, L., & Mitchell, A. (1983). The roles of self-esteem and self-consciousness in the Wortman-Brehm model of reactance and learned helplessness. *Journal of personality and social psychology, 45*, 199.

Brown, J. D., & Cai, H. (2010). Thinking and feeling in the People's Republic of China: Testing the generality of the "laws of emotion". *International Journal of Psychology, 45*, 111-121.

Brown, J. D., & Dutton, K. A. (1995). The thrill of victory, the complexity of defeat: self-esteem and people's emotional reactions to success and failure. *Journal of personality and social psychology, 68*, 712.

Brown, J. D., & Marshall, M. A. (2001). Self-esteem and emotion: Some thoughts about feelings. *Personality and Social Psychology Bulletin, 27*, 575-584.

Bulik, C. M., Carpenter, L. L., Kupfer, D. J., & Frank, E. (1990). Features associated with suicide attempts in recurrent major depression. *Journal of Affective Disorders, 18*, 29-37.

Chung, J. M., Robins, R. W., Trzesniewski, K. H., Noftle, E. E., Roberts, B. W., & Widaman, K. F. (2014). Continuity and change in self-esteem during emerging adulthood. *Journal of personality and social psychology, 106*, 469.

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology, 13*, 99-125.

CRD. (2009). *Systematic Reviews: CRD's guidance for undertaking reviews in health care*: Centre for Reviews and Dissemination University of York.

Dalal, A. K., & Sethi, A. (1988). An attributional study of high and low need achievers in India. *The Journal of social psychology, 128*, 55-64.

Dray, J., Bowman, J., Freund, M., Campbell, E., Wolfenden, L., Hodder, R. K., & Wiggers, J. (2014). Improving adolescent mental health and resilience through a resilience-based intervention in schools: study protocol for a randomised controlled trial. *Trials, 15*, 1.

Dutton, K. A., & Brown, J. D. (1997). Global self-esteem and specific self-views as determinants of people's reactions to success and failure. *Journal of personality and social psychology, 73*, 139.

Follette, V. M., & Jacobson, N. S. (1987). Importance of attributions as a predictor of how people cope with failure. *Journal of personality and social psychology, 52*, 1205.

Forsyth, D. R., & McMillan, J. H. (1981). Attributions, affect, and expectations: A test of Weiner's three-dimensional model. *Journal of Educational Psychology, 73*, 393.

Frost, R. O., Turcotte, T. A., Heimberg, R. G., Mattia, J. I., Holt, C. S., & Hope, D. A. (1995). Reactions to mistakes among subjects high and low in perfectionistic concern over mistakes. *Cognitive Therapy and Research, 19*, 195-205.

Goldhagen, B. E., Kingsolver, K., Stinnett, S. S., & Rosdahl, J. A. (2015). Stress and burnout in residents: impact of mindfulness-based resilience training. *Advances in medical education and practice, 6*, 525.

Green, B. F., & Hall, J. A. (1984). Quantitative methods for literature reviews. *Annual review of psychology, 35*, 37-54.

Hall, L. H., Johnson, J., Watt, I., Tsipa, A., & O'Connor, D. B. (2016). Healthcare Staff Wellbeing, Burnout, and Patient Safety: A Systematic Review. *PloS One, 11*, e0159015.

Higgins, J. P., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions* (Vol. 5): Wiley Online Library.

Hill, A. P., Hall, H. K., Duda, J. L., & Appleton, P. R. (2011). The cognitive, affective and behavioural responses of self-oriented perfectionists following successive failure on a muscular endurance task. *International Journal of Sport and Exercise Psychology, 9*, 189-207.

Hovey, J. D. (1999). Moderating influence of social support on suicidal ideation in a sample of Mexican immigrants. *Psychological Reports, 85*, 78-79.

Ingram, R. E., Johnson, B. R., Bernet, C. Z., Dombeck, M., & Rowe, M. K. (1992). Vulnerability to distress: Cognitive and emotional reactivity in chronically self-focused individuals. *Cognitive Therapy and Research, 16*, 451-472.

Johnson, J. (2016). Resilience: The Bi-dimensional Framework. . In A. M. Wood & J. Johnson (Eds.), *Positive Clinical Psychology*. Chichester: Wiley.

Johnson, J., Gooding, P., & Tarrier, N. (2008a). Suicide risk in schizophrenia: Explanatory models and clinical implications, The Schematic Appraisal Model of Suicide (SAMS). *Psychology and Psychotherapy: Theory, Research and Practice, 81*, 55-77.

Johnson, J., Gooding, P. A., Wood, A. M., & Tarrier, N. (2010a). Resilience as positive coping appraisals: Testing the schematic appraisals model of suicide (SAMS). *Behaviour Research and Therapy, 48*, 179-186.

Johnson, J., Gooding, P. A., Wood, A. M., Taylor, P. J., Pratt, D., & Tarrier, N. (2010b). Resilience to suicidal ideation in psychosis: Positive self-appraisals buffer the impact of hopelessness. *Behaviour Research and Therapy, 48*, 883-889.

Johnson, J., Gooding, P. A., Wood, A. M., Taylor, P. J., & Tarrier, N. (2011a). Trait reappraisal amplifies subjective defeat, sadness, and negative affect in response to failure versus success in nonclinical and psychosis populations. *Journal of Abnormal Psychology, 120*, 922-934.

Johnson, J., Jones, C., Lin, A., Wood, S., Heinze, K., & Jackson, C. (2014). Shame amplifies the association between stressful life events and paranoia amongst young adults using mental health services: Implications for understanding risk and psychological resilience. *Psychiatry research, 220*, 217-225.

Johnson, J., O'Connor, D., Jones, C., Jackson, C., Hughes, G., & Ferguson, E. (In press). Reappraisal Predicts A Weaker Association Between Stress and Negative Mood Measured Over 14 Days: Implications for Understanding Psychological Resilience. *European Journal of Personality*.

Johnson, J., Tarrier, N., & Gooding, P. (2008b). An investigation of aspects of the cry of pain model of suicide risk: the role of defeat in impairing memory. *Behaviour Research and Therapy, 46*, 968-975.

Johnson, J., & Wood, A. M. (2016). Integrating Positive and Clinical Psychology: Viewing Human Functioning as Continua from Positive to Negative Can Benefit Clinical Assessment, Interventions and Understandings of Resilience. *Cognitive Therapy and Research*.

Johnson, J., Wood, A. M., Gooding, P., Taylor, P. J., & Tarrier, N. (2011b). Resilience to suicidality: The buffering hypothesis. *Clinical Psychology Review, 31*, 563-591.

Jones, M. V., & Sheffield, D. (2007). The impact of game outcome on the well ‐ being of athletes. *International Journal of Sport and Exercise Psychology, 5*, 54-65.

Jones, N. P., Papadakis, A. A., Orr, C. A., & Strauman, T. J. (2013). Cognitive processes in response to goal failure: A study of ruminative thought and its affective consequences. *Journal of social and clinical psychology, 32*.

Karabenick, S. A., & Marshall, J. M. (1974). Performance of females as a function of fear of success, fear of failure, type of opponent, and performance ‐ contingent feedback1. *Journal of Personality, 42*, 220-237.

Keeley, M. L., Storch, E. A., Merlo, L. J., & Geffken, G. R. (2008). Clinical predictors of response to cognitive-behavioral therapy for obsessive–compulsive disorder. *Clinical psychology review, 28*, 118-130.

Kernis, M. H., Brockner, J., & Frankel, B. S. (1989). Self-esteem and reactions to failure: The mediating role of overgeneralization. *Journal of Personality and Social Psychology, 57*, 707.

Knopp, J., Knowles, S., Bee, P., Lovell, K., & Bower, P. (2013). A systematic review of predictors and moderators of response to psychological therapies in OCD: do we have enough empirical evidence to target treatment? *Clinical psychology review, 33*, 1067-1081.

Light, R., & Smith, P. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review, 41*, 429-471.

Lynch, K. B., Geller, S. R., & Schmidt, M. G. (2004). Multi-year evaluation of the effectiveness of a resilience-based prevention program for young children. *Journal of Primary Prevention, 24*, 335-353.

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. *Research synthesis methods, 5*, 221-234.

Martin-Krumm, C. P., Sarrazin, P. G., Peterson, C., & Famose, J.-P. (2003). Explanatory style and resilience after sports failure. *Personality and individual differences, 35*, 1685-1695.

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist, 56*, 227-238.

Masten, A. S., & Powell, J. L. (2003). A resilience framework for research, policy, and practice. In S. S. Luthar (Ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 1-25). New York, US: Cambridge University, US.

McCarty, C. A., Mason, W. A., Kosterman, R., Hawkins, J. D., Lengua, L. J., & McCauley, E. (2008). Adolescent school failure predicts later depression among girls. *Journal of Adolescent Health, 43*, 180-187.

McManus, S., Bebbington, P., Jenkins, R., & Brugha, T. (2016). Mental Health and Wellbeing in England Adult Psychiatric Morbidity Survey 2014. Leeds: NHS Digital.

Mealer, M., Conrad, D., Evans, J., Jooste, K., Solyntjes, J., Rothbaum, B., & Moss, M. (2014). Feasibility and acceptability of a resilience training program for intensive care unit nurses. *American Journal of Critical Care, 23*, e97-e105.

Mednick, S. (1962). The associative basis of the creative process. *Psychological review, 69*, 220.

Mendelson, T., & Gruen, R. J. (2005). Self-criticism, failure, and depressive affect: A test of personality–event congruence and symptom specificity. *Cognitive therapy and research, 29*, 301-314.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151*, 264-269.

Morris, M., & Tiggemann, M. (1999). Depressive reactions to academic failure: A test of the diathesis-stress and mediation components of the hopelessness model of depression. *Australian journal of psychology, 51*, 98-103.

Niiya, Y., & Crocker, J. (2008). Mastery goals and contingent self-worth: A field study. *Revue internationale de psychologie sociale, 21*, 135-154.

Niiya, Y., Crocker, J., & Bartmess, E. N. (2004). From Vulnerability to Resilience Learning Orientations Buffer Contingent Self-Esteem From Failure. *Psychological Science, 15*, 801-805.

Nummenmaa, L., & Niemi, P. (2004). Inducing affective states with success-failure manipulations: a meta-analysis. *Emotion, 4*, 207.

Park, L. E., Crocker, J., & Kiefer, A. K. (2007). Contingencies of self-worth, academic failure, and goal pursuit. *Personality and Social Psychology Bulletin*.

Peterson, C., Semmel, A., von Baeyer, C., Abramson, L., Y., Metalsky, G., I., & Seligman, M., E., P.,. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research, 6*, 287-300.

Priester, M. J., & Clum, G. A. (1992). Attributional style as a diathesis in predicting depression, hopelessness, and suicide ideation in college students. *Journal of Psychopathology and Behavioral Assessment, 14*, 111-122.

Priester, M. J., & Clum, G. A. (1993). Perceived problem-solving ability as a predictor of depression, hopelessness, and suicide ideation in a college population. *Journal of Counseling Psychology, 40*, 79-85.

Putwain, D. (2008). *Examination stress and test anxiety*.

Reinherz, H. Z., Giaconia, R. M., Hauf, A. M. C., Wasserman, M. S., & Silverman, A. B. (1999). Major depression in the transition to adulthood: risks and impairments. *Journal of abnormal psychology, 108*, 500.

Ridner, S. H. (2004). Psychological distress: concept analysis. *Journal of advanced nursing, 45*, 536-545.

Riketta, M., & Ziegler, R. (2007). Self ‐ ambivalence and reactions to success versus failure. *European Journal of Social Psychology, 37*, 547-560.

Rosenberg, M. (1965). *Society and the adolescent self-Image*. Princeton, NJ: Princeton University Press.

Saltzman, W. R., Lester, P., Beardslee, W. R., Layne, C. M., Woodward, K., & Nash, W. P. (2011). Mechanisms of risk and resilience in military families: Theoretical and empirical basis of a family-focused resilience enhancement program. *Clinical child and family psychology review, 14*, 213-230.

Sanna, L. J. (1996). Defensive pessimism, optimism, and stimulating alternatives: Some ups and downs of prefactual and counterfactual thinking. *Journal of personality and social psychology, 71*, 1020.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*: Sage publications.

Sellers, S. L., Neighbors, H. W., & Bonham, V. L. (2011). Goal ‐ striving stress and the mental health of college ‐ educated Black American Men: The protective effects of system ‐ blame. *American Journal of Orthopsychiatry, 81*, 507.

Shalon, M., & Strube, M. J. (1988). Type A behavior and emotional responses to uncertainty: A test of the self-appraisal model. *Motivation and emotion, 12*, 385-398.

Sirriyeh, R., Lawton, R., Gardner, P., & Armitage, G. (2010). Coping with medical error: a systematic review of papers to assess the effects of involvement in medical errors on

healthcare professionals' psychological well-being. *Quality and Safety in Health Care, 19*, 1-8.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y).* . Palo Alto, California, USA: Consulting Psychologists Press.

Stiensmeier ‐ Pelster, J. (1989). Attributional style and depressive mood reactions. *Journal of Personality, 57*, 581-599.

Stoeber, J., Schneider, N., Hussain, R., & Matthews, K. (2014). Perfectionism and Negative Affect After Repeated Failure. *Journal of Individual Differences*.

Sweeney, P. D., & Wells, L. E. (1990). Reactions to Feedback about Performance: A Test of Three Competing Models1. *Journal of Applied Social Psychology, 20*, 818-834.

Thompson, T., & Dinnel, D. L. (2007). Is Self ‐ worth Protection Best Regarded as Intentional Self ‐ handicapping Behaviour or an Outcome of Choking Under Pressure? *Educational Psychology, 27*, 509-531.

Thompson, T., Foreman, P., & Martin, F. (2000). Impostor fears and perfectionistic concern over mistakes. *Personality and Individual Differences, 29*, 629-647.

West, C. P., Tan, A. D., Habermann, T. M., Sloan, J. A., & Shanafelt, T. D. (2009). Association of resident fatigue and distress with perceived medical errors. *JAMA, 302*, 1294-1300.

Williams, J. M. G. (1997). *Cry of Pain*. London: Penguin.

Williams, J. M. G., Crane, C., Barnhofer, T., Brennan, K., Duggan, D. S., Fennell, M. J., Hackmann, A., Krusche, A., Muse, K., & Von Rohr, I. R. (2014). Mindfulness-based cognitive therapy for preventing relapse in recurrent depression: a randomized dismantling trial. *Journal of Consulting and Clinical Psychology, 82*, 275.

Wilson, G. V., & Kerr, J. H. (1999). Affective responses to success and failure:: a study of winning and losing in competitive rugby. *Personality and Individual Differences, 27*, 85-99.

Woo, T., & Mix, P. (1997). Self-enhancing reactions to performance feedback in an academic setting. *Journal of Social Behavior and Personality, 12*, 481.

Wytykowska, A., & Gabińska, A. (2015). The effect of emotions, promotion vs. prevention focus, and feedback on cognitive engagement. *Polish Psychological Bulletin, 46*, 350-361.

Zuckerman, M., & Lubin, B. (1965). *Manual for the Multiple Affect Check List*. San Diego, USA

Educational and Industrial Testing Service.

Table 1

Characteristics of included experimental studies

Author/ year	Country	Study design	Resilience variable	Outcome variable/s	Failure manipulation	Significant interactions	Pattern of the interaction	Participant sample	Sample size	Men (%)	Age *M*
Angoli et al. (2015; Study 1)	Italy	Experimental	Emotional Intelligence (Emotional Intelligence Questionnaire–Short Form (Petrides & Furnham, 2006)	30 items from the PANAS-X (Watson & Clark, 1990), measuring sadness, guilt, fatigue, joviality and self-assurance affect	False feedback - positive or negative feedback on a computerised task. Task involved helping a child	Emotional intelligence interacted with failure to predict sadness and guilt	Positive feedback predicted reduction in sadness and guilt in low Emotional Intelligence but not high Emotional Intelligence individuals	Undergraduates	63	55.6	24.1
Angoli et al. (2015; Study 2)	Italy	Experimental	Emotional Intelligence (Emotional Intelligence Questionnaire–Short Form (Petrides & Furnham, 2006)	30 items from the PANAS-X (Watson & Clark, 1990), measuring sadness, guilt, fatigue, joviality and self-assurance affect	False feedback - positive or negative feedback on a computerised task. Task did not involve helping another person	None	Not applicable - there were no significant interactions	Undergraduates	59	53.3	24.52
Anshel & Mansouri (2005)	USA	Experimental	Perfectionism (Organisation subscale of The Multiple Perfectionism Scale; Frost et al., 1990)	Negative and positive affect (Children's Arousal Scale – Adult version; Anshel & Martin, 1996)	No feedback (control condition) or false failure feedback (experimental condition) on a body-balancing task on a stabilometer for 20 trials	None	Not applicable - there were no significant interactions	College-aged male athletes	30	100	Mean age not available. Range: 19.6-22.8
Basgall & Snyder (1988)	USA	Experimental	Locus of Control (Internal-External Locus of Control Scale; Nowicki & Duke, 1974)	Anxiety, depression and hostility (Multiple Affect Adjective Checklist; Zuckerman & Lubin, 1965)	False success or failure feedback on a purported test of social perceptiveness.	Locus of control interacted with failure	Individuals with external locus of control became more depressed after failure than individuals with internal locus of control	Undergraduates scoring in the upper and lower quartiles on Locus of Control from an initial sample of 600	96	0	Not available
Besser et al. (2004)	Israel	Experimental	Self-Oriented Perfectionism and Socially-Prescribed Perfectionism (Multidimensional Perfectionism	Positive affect, dysphoria, hostility and anxiety measured using visual analogue scales of 18 mood	False success or failure feedback on a computerised task.	Self-Oriented Perfectionism interacted with feedback to predict positive affect	Under negative feedback, high self-oriented perfectionists reported a decrease in post-task positive affect. When the	Undergraduates	100	50	21.75

			Scale; Hewitt & Flett 1991).	adjectives.			feedback was positive, high self-oriented perfectionists reported a significant increase in positive affect				
Besser et al. (2008)	Israel	Experimental	Self-Oriented Perfectionism and Socially-Prescribed Perfectionism (Multidimensional Perfectionism Scale; Hewitt & Flett 1991).	Positive affect, dysphoria, hostility and anxiety measured using visual analogue scales of 18 mood adjectives; Performance self-esteem and social self-esteem (modified version of the Current Thoughts Scale ; Heatherton & Polivy, 1991)	1) False feedback - positive or negative feedback on a computerised task, and 2) Objective errors/mistakes	Socially prescribed perfectionism moderated the impact of objective performance on dysphoria and positive affect, and the impact of feedback on positive affect and performance self esteem.	High socially prescribed perfectionism was associated with 1) low post-task performance self-esteem and this was stronger under negative feedback, 2) increased dysphoria and reductions in positive affect when there were higher levels of objective errors, 3) decreases in positive affect in response to negative feedback	Undergraduates	200	50	23.63
Bodroza (2011)	Serbia	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-esteem (global self-esteem scale; Opacic & Bodroza, in preparation at the time of publication)	Depression, anxiety and anger (Pofile of affective states; Popov, 2007).	False success or failure feedback on a computerised task.	None	Not applicable - there were no significant interactions	Undergraduates	90	0	21.25
Brockner	USA	Experimental	Self esteem	Confident, upset,	Insoluble	None	Not applicable -	Undergraduates	78	33	Not

(1983; Study 1)		. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	(revised Janis-Field Scale; Eagly 1967) and self-consciousness (Private self-consciousness subscale of the Self-Consciousness Scale; Fenigstein, Scheier & Buss, 1975)	frustrated, angry, and depressed, measured using a 41-item measure	anagrams task (control condition v failure)		there were no significant interactions				available
Brockner (1983; Study 2)	USA	Experimental . However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self esteem (revised Janis-Field Scale; Eagly 1967) and self-consciousness (Private self-consciousness subscale of the Self-Consciousness Scale; Fenigstein, Scheier & Buss, 1975)	Confident, upset, frustrated, angry, and depressed, measured using a 41-item measure	Insoluble anagrams task (control condition v small failure v extended failure)	None	Not applicable - there were no significant interactions	Undergraduates	119	52	Not available
Brown & Cai (2010; Study 1)	USA (but included Chinese participants only)	Experimental . However, baseline scores were not recorded, and as such, post-experimental findings could be	Attributional style - single item measuring the extent to which participants thought their performance was due to their integrative orientation ability[b]	Self relevant emotions (proud, pleased with myself, ashamed, humiliated, e.g., Brown & Dutton, 1995). Some validation information provided.	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	None, but there was a trend towards an interaction between attributional style moderating the association between	No significant interactions, however, there was a trend. In the failure condition, both high and low ability attribution individuals report the same levels of self-worth, but in the	Undergraduates	55	25	19.46

		explained by baseline differences. These studies are susceptible to selective reporting.				success/failure and feelings of self worth (p=.065).	success condition, high ability attribution individuals report higher levels of self-worth				
Brown & Cai (2010; Study 2)	USA (but included American and Chinese participants)	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Attributional style - single item measuring the extent to which participants thought their performance was due to their integrative orientation ability[b]	Self relevant emotions (proud, pleased with myself, ashamed, humiliated, e.g., Brown & Dutton, 1995). Some validation information provided.	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	Attributional style moderated associations between success/failure and feelings of self worth	Cross-over effect - those with high ability attribution showed higher feelings of self-worth in the success condition, but lower feelings of self-worth in the failure condition	Undergraduates	310 (144 Chinese)	29	Not available
Brown & Dutton (1995; Study 1)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965)	8-item emotion scale. The scale consisted of two subscales: (1) outcome-dependent emotion (glad, happy, sad, unhappy) and (2) self relevant emotions (proud, pleased with myself, ashamed, humiliated).	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	Self-esteem interacted with failure to predict levels of self relevant emotions	High self-esteem buffers individuals from reduced positive emotion in the face of failure	Undergraduates	172	23	Not available

Brown & Dutton (1995; Study 2)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965)	8-item emotion scale. The scale consisted of two subscales: (1) outcome-dependent emotion (glad, happy, sad, unhappy) and (2) self relevant emotions (proud, pleased with myself, ashamed, humiliated).	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	Self-esteem interacted with failure to predict levels of self relevant emotions	High self-esteem buffers individuals from reduced positive emotion in the face of failure	Undergraduates	129	39	Not available
Brown & Marshall (2001; Study 2)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-esteem measured with i) Self-Esteem Questionnaire (Rosenberg, 1965), and ii) Texas Social Behavior Inventory (Helmreich & Stapp, 1974)	Self relevant emotions (proud, pleased with myself, ashamed, humiliated, e.g., Brown & Dutton, 1995).	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	Self-esteem measured using both the SEQ and the TSBI interacted with failure to predict emotion	High self-esteem buffered the association between failure and higher levels of negative self-relevant emotions	Undergraduates	291	32	Not available
Brown & Marshall (2001; Study 3)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965)	Self-relevant emotion scale formed from four items (proud, pleased with myself, humiliated, ashamed, e.g., Brown & Dutton,	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	Self-esteem interacted with failure to predict self-relevant emotions	High self-esteem buffered the association between failure and higher levels of negative self-relevant emotions	Undergraduates	72	32	Not available

		could be explained by baseline differences. These studies are susceptible to selective reporting.		1995); Non-self-relevant emotions measured using 18 items from the Positive and Negative Affect Scales (the total scale minus “proud” and “ashamed”; PANAS; Watson 1988).							
Dalal & Sethi (1988)	India	Experimental	Need for achievement (Indian version of the Edwards Personality Preference Schedule; Dhavan 1982)	Single mood scale measuring positive-negative affect (created from 10 bipolar emotion-related adjectives responded to on 7-point scales)	Anagrams task. Success or failure manipulated by the giving of easy (success condition) or difficult (failure condition) tasks	None	Not applicable - there were no significant interactions	Undergraduates	48	Not available	Not available
Dutton & Brown (1997; Study 1)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965)	Self relevant emotions (proud, pleased with myself, ashamed, humiliated, e.g., Brown & Dutton, 1995). Some validation information provided.	False success or failure feedback on a computerised task (Remote Associates Test; Mednick, 1962)	Self-esteem interacted with failure to predict emotion	Plot indicates that high self-esteem buffers individuals from experiencing negative emotions in the face of failure	Undergraduates	191	33	Not available
Dutton & Brown (1997; Study 2)	USA	Experimental. However, baseline scores were not recorded, and as such,	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965) and a single composite	Self relevant emotions (proud, pleased with myself, ashamed, humiliated, e.g., Brown & Dutton,	False success or failure feedback on a computerised task (Remote Associates Test;	Both measures of self-esteem interacted with failure to predict emotions	Plots indicate that high self-esteem buffers individuals from experiencing negative emotions in the face of failure	Undergraduates	136	31	Not available

		post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	measure based on how well participants thought 10 positive and negative attributes described them (e.g., intelligent, athletic, attractive, uncoordinated, unattractive, inconsiderate).	1995). Some validation information provided.	Mednick, 1962)						
Frost et al. (1995)	USA	Experimental, but it is unclear whether baseline affect was controlled for in the analysis	Concern Over Mistakes (CM) subscale of the Multidimensional Perfectionism Scale (Frost, 1990)	Negative affect (measure not clearly defined in paper)	Number of mistakes in a computerised task, high mistakes v low mistakes	Concern over mistakes interacted with number of mistakes to predict negative affect	Low perfectionism buffers the impact of being in the high-mistake task on low mood	Undergraduates	64	Not available	Not available
Hill et al. (2011)	UK	Experimental, but all participants received the failure condition, and their scores were compared to their own baseline scores.	Self-oriented perfectionism subscale of the Multidimensional Perfectionism Scale (Hewitt and Flett 1991)	Positive and Negative Affect measured using the Positive and Negative Affect Scales (PANAS; Watson 1988)	Performance feedback on a cycling task manipulated to ensure failure to meet personal goals. All participants received the failure induction, scores on outcome measure compared pre and post	None	Not applicable - there were no significant interactions	Undergraduates	68	71	19.75
Ingram et al. (1992; Study 1)	USA	Experimental	Private self-consciousness measured using 10 items from the Self-Consciousness Scale (Fenigstein, 1975)	Multiple Affect Adjective Checklist (MAACL; Zuckerman & Lubin, 1965). Comprises three subscales: Anxiety, depression and	False failure or success feedback on a bogus intelligence paper-and-pencil test	None, although there were trends towards self-consciousness interacting with failure to predict the overall mood score and	Not applicable - there were no significant interactions	Undergraduates	58	Not available	Not available

				hostility. Overall score and the three subscales were investigated		depression.					
Johnson et al. (2011a; Study 1)	UK	Experimental	Trait Suppression and Trait Reappraisal measured using the Emotion Regulation Questionnaire (Gross & John 2003)	Positive and Negative Affect measured using the Positive and Negative Affect Scales (PANAS; Watson 1988), and Visual Analogue Scales (VAS) of five mood states (defeat, sadness, calmness, happiness, and frustration)	False success or failure feedback on a task (Remote Associates Test; Mednick, 1962)	Trait reappraisal interacted with failure to predict negative affect on the PANAS and VAS scales of defeat, sadness and calmness	Low levels of trait reappraisal buffer the association between failure and higher negative mood, and amplify feelings of calmness in the face of failure	Undergraduates	120	23	20.53
Johnson et al. (2011a; Study 2)	UK	Experimental	Trait Suppression and Trait Reappraisal measured using the Emotion Regulation Questionnaire (Gross & John 2003)	Visual Analogue Scales of five mood states (defeated, sad, calm, happy, and frustrated)	False success or failure feedback on a task (Remote Associates Test; Mednick, 1962)	Trait reappraisal interacted with failure to predict defeat	Low levels of trait reappraisal buffer the association between failure and feelings of defeat	Adults with a diagnosis of a schizophrenia-spectrum disorder	77	77	42.3
Jones et al. (2013)	USA	Experimental	Chronic promotion failure measured using the Computerized Selves Questionnaire (CS; Jones et al., 2009). This measures the discrepancy between participants' goals for themselves and where they perceive themselves to be	Dejection and Quiescence measured using items from the Sadness and Serenity subscales of Positive and Negative Affect Scale – Expanded Form (PANAS-X; Watson & Clark, 1990)	Writing task to elicit memories of 'promotion failure', 'prevention failure' or control memories	Chronic promotion failure interacted with failure condition to predict dejection	Low levels of chronic promotion failure buffer the impact of failure memories on dejection	Undergraduates	78	21	26.37
Karabenic	USA	Experimental	Projective	7-point bipolar	False failure or	None	Not applicable -	Undergraduates	252	0	Not

k & Marshall (1974)	. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting	measure of fear of success using fear of success stories (Horner 1968); Fear of failure measured using the Debilitating Anxiety Scale of the Achievement Anxiety Test (Alpert & Haber, 1960)	emotion scales of depression-pleasure; unembarassment-embarrassment; luck-skill; happy-unhappy; uncomfortable-comfortable; superior-inferior; relaxed-nervous	success or neutral [equal] feedback on a paper task. Some participants were compared directly to a confederate opponent, others to normed scores		there were no significant interactions				available
Mendelson & Gruen (2005)	USA	Experimental. Mood change was measured immediately following the failure induction and again 24 hours later	Self-criticism and dependency (Depressive Experiences Questionnaire; Blatt et al., 1976)	Three types of depressive affect: Introjective and anaclitic depressive affect (Emotion Questionnaire, Zuroff & Mongrain, 1987) and Depression–Dejection (subscale from the Profile of Mood States, McNair, Lorr, & Droppleman, 1971)	In the failure condition, false feedback was provided in response to a version of the Ravens Progressive Matrices (Raven, Court & Raven, 1985). In the control condition, participants sat quietly with a book of nature pictures	Self-criticism interacted with failure to predict changes in introjective depressive affect immediately following the failure. Self-criticism and dependency interacted with failure to predict anaclitic depression immediately following the failure	Pattern of the interactions not displayed or described	Undergraduates	125	36.8	19.42
Niiya et al. (2004)	USA	Experimental	Academic subscale of the Contingencies of Self-Worth Scale (Crocker, et al., 2003)	State self-esteem (Heatherton & Polivy, 1991), comprising three correlated factors: performance, social, and	False success (i.e., a score of 97th percentile) or failure (i.e., a score of 45th percentile) feedback on a	None	Not applicable - there were no significant interactions	Undergraduates	128	26.6	Not available

			appearance state self-esteem	Graduate Record Examination (GRE) test composed of reading comprehension, quantitative comprehension, and analytical reasoning questions						
Park et al. (2007; Study 1) [a]	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965); Academic subscale of the Contingencies of Self-Worth Scale (Crocker, et al., 2003)	State self-esteem adapted from the Rosenberg Self-Esteem Questionnaire to measure feelings at that moment; Positive and negative affect measured using 7-point rating scales for positive affect items (e.g., happy, cheerful; 7 items) and negative affect (e.g., angry, depressed; 7 items)	Remotes Associates Test (McFarlin & Blascovich, 1984). Participants in the control condition rated words for their favourite, and were given no evaluative feedback	None	Not applicable - there were no significant interactions	Undergraduates	122	35.2 19.01
Park et al. (2007; Study 2) [a]	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965); Academic subscale of the Contingencies of Self-Worth Scale (Crocker, et al., 2003)	Implicit affect measured using the IAT (Greenwald et al., 1998), a computerized reaction time task that measures the relative speed of associations made between target concepts and attributes. Participants categorized words	Remotes Associates Test (McFarlin & Blascovich, 1984). Participants in the non-failure condition completed an easy version of the test which ensured success	None	Not applicable - there were no significant interactions	Undergraduates	109	53.2 19.79

		susceptible to selective reporting		related to the self and other with words related to failure (e.g., worthless, failure, incompetent) and words related to success (e.g., worthy, success, competent)							
Riketta & Ziegler (2007)	Germany	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting	Experienced ambivalence (e.g., “I have positive and negative feelings toward myself at the same time”; Riketta & Ziegler, 2005); Structural ambivalence (e.g., “please consider only the positive (negative) aspects of yourself-image.. how positive do you find yourself?”; Thompson et al., 1995); Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965)	Two outcomes. The first was self-feeling items of proud, ashamed, humiliated and satisfied and mood items of depression, good-humour, sad and happy (Brown & Dutton, 1995). The second was state self-esteem (Heatherton & Polivy, 1991)	Computerised task fixed to produce success (easy version) or failure (hard version). Based on the Ravens Advanced Progressive Matrices (APM), a standardized nonverbal intelligence test	Four hierarchical regression analyses tested each type of ambivalence separately in relation to the two outcomes. Of those testing structural ambivalence, structural ambivalence interacted with failure to predict state self-esteem. Self-esteem interacted with failure to predict self-feelings and mood. Of those testing experienced ambivalence, self-esteem interacted with failure to predict self-feelings and mood.	Low structural ambivalence buffered against the negative impact of failure upon state self-esteem. High self-esteem buffered participants from a drop in state self esteem in response to failure.	Undergraduates	87	54	21.84

Sanna (1996; Study 4)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting	Defensive pessimism/optimism (the propensity to use defensive pessimistic or optimistic strategies in academic achievement situations; Norem & Illingworth, 1993). Participants scoring in the upper third were classed as "optimists" and in the lower third, as "pessimists". Participants were selected from a larger group of 454 for scoring high or low on this scale	Participants indicated the extent to which a series of positive and negative adjectives reflected their mood with regard to the upcoming exam. All items were scored to produce a single mood outcome variable	Anagrams task. The same task was given in success and failure conditions but feedback manipulated to tell the participant they had either scored in the top 20th (success condition) or bottom 20th (failure condition) centiles	None	Not applicable - there were no significant interactions	Undergraduates	87	Not available	Not available
Shalon & Strube (1988)	USA	Experimental. However, baseline scores were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Type A/ Type B behaviour pattern measured using the Jenkins Activity Survey Form (Krantz, Glass, & Snyder, 1974. Participants classified as Type A's (scores of 9 or greater) or Type Bs (scores of 8 or less)	Mood scales of anxiety, nervousness, frustration, anger, and depression	Anagrams task (success v failure). In the success condition, participants completed easy anagrams and were told that their score was better than; or equal to, 78%of students. In the failure condition, participants completed a very difficult set of anagrams and were told that 42% of the people taking the test did	None, although there was a trend towards Type A/Type B behaviour moderating the association between failure and anxiety ($p<0.55$).	Not available	Undergraduates	80	50	Not available

				better than them						
Steinsmeier-Pelster (1989)	Germany	Experimental	Attributional style was assessed with the negative items from the German Attributional Style Questionnaire (GASQ, Stiensmeier et al , 1985), based on the original ASQ (Peterson et al., 1982)	Mood index was created by totalling the Carefreeness (reverse-scored), Happiness (reverse-scored), and Depression scales from the Mehrdimensionale Stimmungsfragebogen (Hecheltjen & Mertesdorf, 1973)	Participants completed the task together with a confederate. Two versions of the Raven Progressive Matrices (Raven, 1974/1975) were used. The difficulty level of the tasks and the behavior of the confederate were manipulated to induce failure and success. False feedback not given	Failure interacted with attributional style to predict mood	Negative attributional style amplified negative mood in the failure condition only.	Undergraduates	46	0	20.4
Stoeber et al. (2014)	UK	Experimental , but mood was not recorded at baseline. As such, for the interactions testing mood after the first task, findings could be explained by baseline differences. For analyses of mood after the second task, prior mood was included as a control variable	Self-oriented perfectionism and socially prescribed perfectionism (Multidimensional Perfectionism Scale; Hewitt & Flett 2004)	Three mood measures. Anxiety (a short form of the State–Trait Anxiety Inventory; Spielberger,Gorsuch, Lushene, Vagg, & Jacobs, 1983); depression (subscale from a short form of the Profile of Mood States, McNair, Lorr, & Droppleman, 1971); anger (Feeling Angry subscale of the State–Trait Anger Expression Inventory; Spielberger, 1999). Mood measured after	False feedback to induce success and failure provided in response to computerised tasks involving identifying whether pictures of rotated figures were the same figure. Each participant completed two similar tasks and mood was measured after each task	Socially prescribed perfectionism interacted with failure to predict anxiety, depression and anger after the first task. Socially prescribed perfectionism interacted with failure to predict anger after the second task and self-oriented perfectionism interacted with failure to predict anxiety after the second task	Socially prescribed perfectionism amplified the association between failure and anxiety, depression and anger after the first task. Socially prescribed perfectionism amplified the association between failure and anger after the second task, and self-oriented perfectionism amplified the association between failure and anxiety after the second task.	Undergraduates	100	50	21.35

			the first failure and again after the second failure							
Thompson & Dinnel (2007)	Australian	Experimental. However, baseline scores of dependent variables were not recorded, and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-worth protection (the extent to which participants want to avoid failure measured using the Self-Worth Protection Scale; Thompson & Dinnel, 2003). Participants were selected from a larger group of 235 for scoring high (upper third) or low (lower third) on this scale.	Negative affect index created using three items (guilt, shame, humiliation) from the PANAS (Watson, Clark, & Tellegen, 1988)	Three conditions, success, face-saving failure (where participants were informed that ability on the task had not been found to be a particularly good indicator of overall ability) and humiliating failure (where participants were informed that ability on the task was a reliable indicator of general intelligence). Task was a computer discrimination task. In the failure conditions, false failure feedback was given. In the success condition, feedback was related to performance	Self-worth protection interacted with performance feedback condition to predict negative affect.	Students high in self-worth protection reported greater negative affect following humiliating failure than students low in self-worth protection, as was the case following success, but not following face-saving failure.	Undergraduates	72	48.6	22.85
Thompson et al. (2000)	Australian	Experimental. However, baseline scores of dependent variables were not recorded (or in the case of STAI anxiety, simply not	Imposter fears (modified version of the Clance Impostor Phenomenon scale; Clance, 1985). Participants who scored as "imposters" or "non-imposters"	Four outcome measures: positive affect, negative affect (PANAS Scales; Watson, Clark & Tellegen, 1988), post-task anxiety (State-Trait anxiety Inventory, Spielberger,T.	Computerised version of the Stroop task. Real feedback given, and incorrect responses emphasised with an "uh oh" sound. Two versions of this: high mistakes	Imposter status interacted with failure to predict single-item anxiety and positive mood	Being a non-imposter buffers against a drop in positive mood/increased anxiety in response to failure	Undergraduates	60	18.3	21

		controlled for), and as such, post-experimental findings could be explained by baseline differences. These studies are susceptible to selective reporting	were drawn from an initial sample of 318 students	Gorsuch, Lushene, Vagg & Jacobs, 1983) and a single-item anxiety measure	frequency and low mistakes frequency. Low mistakes frequency task extremely easy, simply a patch of colour presented on a screen						
Wytykowska & Gobinska (2015)	Poland	Experimental	Promotion vs. prevention orientation (Polish version of Regulatory Focus Questionnaire; Pikuła, 2012). Measures orientations (i.e. anticipatory goal reactions) to new tasks or goals. The higher the score, the more promotion-focused the person is considered to be	Eight emotions were taken into account – feeling depressed, tense, uneasy, discouraged, excited, pleased, interested, and calmness	False feedback on a computerised task. There were three conditions: positive, negative and control. All participants completed a computerised task where they were initially successful. After this, participants completed two further tasks, where they scored roughly the same as the first task (control condition), worse than previously (failure condition) or better than previously (success condition)	Promotion vs. prevention orientation interacted with feedback (failure v success) to predict tension, calmness and feeling pleased	Prevention focus amplified the impact of failure on tension. Pattern of the interaction for calmness and feeling pleased unclear	Senior secondary school students	190	43.16	18.6

[a] This study reported a significant three-way interaction between two potential resilience variables and failure. Please see Supplementary File 2.
[b] In this study, the proposed resilience variable was measured after the experimental induction, which may have introduced bias in responding.

Table 2
Characteristics of included cross-sectional and longitudinal studies

Author/ year	Country	Study design	Resilience variable	Outcome variable/s	Failure experience/ measure	Significant interactions	Pattern of the interaction	Participant sample	Sample size	Men (%)	Age *M*
Abela (2002)[a]	USA	Longitudinal	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965); Inferential style (Cognitive Style Questionnaire; Abramson & Metalsky, 1986)	Residual difference between state depressed mood at baseline and i) on the day of receiving admissions outcome, and ii) four days later (Multiple Affect Adjective Checklist; Zuckerman & Lubin, 1965)	Acceptance or rejection from Penn University	Self-esteem interacted with failure to predict depression four days after receiving admissions outcome	No plot or description of pattern provided	University applicants	136	47.1	Not available
Follette & Jacobson (1987)	USA	Longitudinal	Attributions measured using (1) three subscales of the Expanded Attributional Style Questionnaire (EASQ; Peterson & Seligman, 1984), and (2) the control subscale of the EASQ	Depression subscale of the Multiple Affect Adjective Checklist (MAACL; Zuckerman & Lubin, 1965).	The difference between expected and received university course grade	None	Not applicable - there were no significant interactions	Undergraduates	110	25	Not available
Forsyth & MacMillan (1981)	USA	Cross-sectional	Attributions measured using three items, asking about perceptions of controllability, locus of causality and stability	Visual analogue scales measuring degree to which participants were experiencing 16 mood states	Perceived examination performance	Locus of causality attributions interacted with examination performance to predict overall mood	No plot or description of pattern provided	Undergraduates	223	38	19.3
Kernis et al. (1989)	USA	Longitudinal, but emotion measure only	Tendency to overgeneralize from bad	Participants scored the extent to which they	Examination performance. Participants were	Self-esteem interacted with performance to	High self esteem and low overgeneralization	Undergraduates	149	50	Not available

		completed once at the end of the study. These studies are susceptible to selective reporting	experiences to the overall self-concept (overgeneralizatio n subscale of the Attitudes Toward Self Scale; ATS, Carver & Ganellen, 1983); Self-esteem (Self Esteem Questionnaire; Rosenberg, 1965)	were experiencing 40 specific emotions at that moment. These were factor analysed, and pleasant and unpleasant affect indexes were formed. Unpleasant affect contained 23 words and pleasant affect contained12 words	placed into the high performance group if they had received an A or B grade and their grade was either the same or better than they had expected; they were placed into the low performance group if they had received a C or lower and this was the same or lower than they had expected. Other participants (n = 48) were excluded from the analysis	predict negative affect. Overgeneraliza tion interacted with performance to predict negative affect and positive affect	conferred resilience to higher negative emotion in response to failure, and low overgeneralization conferred resilience to reduced positive affect in response to failure				
Morris & Tiggeman n (1999)	Australia	Longitudinal	Attributional Style Questionnaire (Seligman, Abramson, Semmel, & von Baeyer, 1979). A negative generality score was obtained by averaging the ratings of the globality and stability dimensions. An overall composite was also obtained by averaging all three attributional dimensions	Depressive reaction was assessed by the Beck Depression Inventory - Short Form (Beck, 1967), both immediately following the exam and at the end of the academic year	Subjective performance on an examination (naturally occurring). Calculated as actual grade minus the grade they would be satisfied with (reported before the exam)	Composite attributional style interacted with subjective performance to predict depression immediately following the exam feedback and also at the end of the academic year. Attributional style generality interacted with subjective performance to predict depression at the end of the year.	Pattern of the interaction not plotted. Correlations suggest that attributional styles were only associated with end-of-year depression scores in the failure group	Undergraduates	363	30	22.04
Niiya &	USA	Longitudinal	Academic	Rosenberg Self-	Grade on an	None	Not applicable -	Undergraduates	142	23.9	19.8

Crocker (2008)[a]			subscale of the Contingencies of Self-Worth Scale (Crocker, et al., 2003); Mastery goals subscale of Achievement Goal Scale (Elliot & Church, 1997); Ability-Validation Goal Scale modified from Grant and Dweck (2003), which measures striving to demonstrate or prove ability	Esteem Inventory (Rosenberg, 1965) to which the words “right now” were added to the instructions	assignment which accounted for 15% of the final course grade (naturally occurring)		there were no significant interactions				
Sellers et al. (2011)	USA	Cross-sectional	High-active coping (measured with 12 items, e.g., “I’ve always felt that I could make my life pretty much what I wanted to make of it”).	Mental health was measured using the Mental Health Component Summary of the Medical Outcomes Study Short Form-12 (Ware, Kosinski, & Keller, 1998)	“Goal striving stress" (three items capturing the discrepancy between aspirations and achievement, weighted by the level of disappointment associated with failing to achieve life goals)	None	Not applicable - there were no significant interactions	Black college educated men who were members of a historically black national fraternal organisation	399	100	47.6
Sweeney & Wells (1990)	USA	Longitudinal but mood was not recorded at baseline. As such findings could be explained by baseline differences. These studies are susceptible to selective reporting.	Self-esteem (Self-Esteem Questionnaire; Rosenberg, 1965)	Three measures used to create an “affective index”. 1) single item, “How satisfied were you with the score you received on your exam? (1 = very unsatisfied, 7 = very satisfied).” 2) emotional reaction to the professor “How happy are you with the	Grade on a mid-term college examination	Self-esteem with exam performance to predict affective index scoree	Self-esteem amplified the impact of success/failure on affect	Undergraduates	187	47.1	Not availabl e

				instructor's performance thus far in the term?"(1 = very happy, 2 = pretty happy, 3 = not too happy). 3) Center for Epidemiological Studies, Depression Scale (CES-D, Radloff, 1977)							
Woo & Mix (1997)	USA	Longitudinal but mood was not recorded at baseline. These studies are susceptible to selective reporting	Performance self-esteem (Performance Self-esteem scale; Stake, 1979)	Immediately after exam feedback, positive affect (two items) and negative affect (eight items) was measured	Exam performance. One week prior to the exam, participants indicated their own criteria for "success". Participants whose actual grades equaled or exceeded their criterion performance were classified as the "success" group and those whose grades fell below this were the "failure" group	None	Not applicable - there were no significant interactions	Undergraduates	72	25	Not availabl e

[a] This study reported a significant three-way interaction between two potential resilience variables and failure. Please see Supplementary File 2.

Table 3

Box-score review of interaction effects of proposed resilience variables on the association between failure and emotional distress

Moderator variable	Number of studies	
Academic Self-worth	4	0000
Attributional style	6	++++00
Emotional Intelligence	2	+0
Self-esteem	15	++++++++++00000
Self-oriented perfectionism	4	++00
Socially-prescribed perfectionism	3	++0
Trait Reappraisal	2	++
Trait Suppression	2	00

NB. + = interaction effect significant, 0 = interaction effect significant. Pattern of the interaction not reported here as the complexities of this are beyond the scope of simple symbolic descriptions.