BUFFERING PRECONSCIOUS STRESSOR APPRAISAL: THE PROTECTIVE ROLE OF SELF-EFFICACY

A thesis submitted for the degree of
Doctor of Education (EdD)

By

Timothy William Filtness

School of Sport and Education
Brunel University

July 2013
Buffering Preconscious Stressor Appraisal: the Protective Role of Self-Efficacy

A Thesis presented

by

Timothy William Filtness

Approved as to style and content by:

A

B

C

Professor M Watts, Department Head
School of Sport and Education
Acknowledgements

I would like to thank my outstanding team of supervisors – Dr Dan Bishop, Professor José Gonzalez-Alonso and Professor Mike Watts – without whose support, advice and invaluable guidance, this thesis would never have taken shape.

I would also like to thank my boss, Del Cooke, who not only let me loose in her school during the data collection phase of this study, but who also put up with my absences (both corporeal and cognitive), particularly during SMT meetings. Without your support, Del, this project would still be a dusty item on my ‘to do’ list!

I would like to thank my wife, Claire, who has not only supported me throughout the (long) process of the EdD, but also tolerated my ever-increasing Amazon purchases (sorry about that) and helped keep me pointed in the right direction when I faltered. Without your love and encouragement this project would have not even have made it to the ‘to do’ list – thank you.

Finally, I would like to thank my son, Oscar; who, despite the early mornings, has given everything a sense of meaning.
Abstract

Many cognitive resources contribute towards the appraisal of stressors. Of these, self-efficacy (SE) is widely acknowledged to play a significant role in protecting adolescents from the effects of stress (Bandura, 1997). This study investigated that relationship through the use of a quasi-experimental methodology (Cook & Campbell, 1979) utilising an untreated Control group of 44 adolescent, female participants and an Experimental group of 70 additional participants, all of whom were volunteers drawn from the Sixth Form of a single participating school. The members of both participant groups took part in two rounds of testing, between which the members of the Experimental group were exposed to a significant academic stressor (one or more public A-level examinations). During both test phases, all participants completed the 10-item Perceived Stress Scale self-report (Cohen & Williamson, 1988), the Examination Self-Efficacy Scale instrument (Schwarzer & Jerusalem, 1995) and a bespoke Implicit Association Test (Greenwald et al., 1998) designed to measure implicit stressor appraisal. Significant trends were identified by means of ANCOVA, correlation and regression analyses, and the resulting data were interpreted in terms of a dual process model of stress (Compas, 2004). Results not only concurred with those of previous studies (e.g. Betoret, 2006; Vaezi & Fallah, 2011) by demonstrating a strongly negative correlation between acute academic stress and academic SE, but provided new evidence to suggest that the ‘protective’ effect of SE occurs via a buffering mechanism at the level of preconscious stressor appraisals (Bargh, 1990), which limits the effect of acute stress exposure on preconscious stressor appraisals (e.g. Luecken & Appelhans, 2005).
# Table of Contents

<table>
<thead>
<tr>
<th>List of Figures</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List of Tables</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>xiii</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List of Abbreviations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>xiv</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List of Appendices</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>xv</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 1: Background</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2: Stress</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What is stress?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What are emotions?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emotion regulation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evidence that stress is an emotion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dual processes models of stress</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stress as an emotion - a synopsis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3: Humans as Efficacious Agents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
</tr>
</tbody>
</table>
Introduction 50

The neurological 'self' 52

On reductivism 57

Humans as efficacious agents 64

Efficacy beliefs 68

Summary and research questions 76

Chapter 4: Methodology 78

Choosing the research methodology 78

Choosing the quasi-experimental design 83

Application of the quasi-experimental design 89

Relating the design to the research questions 91

Considerations and Control Procedures 92

The design and validity of the instruments 100

Précis of the methodology 116
Chapter 5: Research Findings

Addressing the research questions

Validity of the instruments

Review of the findings

Chapter 6: Discussion

Individuality of stress

To what extent does academic SE limit the development of acute academic stress?

To what extent does acute academic stress exposure affect the preconscious appraisal of academic stressor stimuli?

To what extent does academic SE affect the preconscious appraisal of academic stressor stimuli?

Next steps

Conclusions

Implications of the conclusions

Limitations of the conclusions
List of Figures

Figure 1  Diagram representing the untreated control group design with one pre-test quasi-experimental methodology (adapted from a similar diagram in Shadish et al., 2002).

Figure 2  Diagram representing the three phases of data collection.

Figure 3  Graphs representing G*Power power analysis for minimum sample size (left, θ = 0.4. right, θ = 0.25).

Figure 4  Average PrS & PoS PSS10 scores for Experimental and Control groups.

Figure 5  Homogeneity of regression for PSS10 data for Control and Experimental groups.

Figure 6  Homogeneity of regression for ESES data for Control and Experimental groups.

Figure 7  Correlation between ESES and ΔPSS10 for Control and Experimental groups.

Figure 8:  Average PrS & PoS SA-IAT scores for Experimental and Control groups.

Figure 9  Homogeneity of regression for SA-IAT data for Control and Experimental groups.
Figure 10  Histograms of SAI-AT Scores showing the difference in distribution between Control (top) and Experimental (bottom) group in PrS (left) and PoS (right) phases.

Figure 11  Correlations between ESES & SAI-AT scores for PrS (left) and PoS (right) phases.

Figure 12  Difference in SE scores between subpopulations of the Experimental group. Left bar represents population with negative ΔSA-IAT scores (PoS-PrS), right bar represents population with positive ΔSA-IAT scores.

Figure 13  Test-retest reliability for the SA-IAT.

Figure 14  Normal Q-Q Plots for PrS SA-IAT data for Experimental group.

Figure 15  Post-stress PSS10 scores for Control group (above) & Experimental group (below).

Figure 16  Distribution of variation in SA-IAT scores between test phases (PoS-PrS) for Control group (above) & Experimental group (below).

Figure A1  “Phases and Sub-processes of self-regulation” adapted from a similar diagram in Zimmerman (2002).
Figure A2: Schematic illustration of the Stressor Appraisal Implicit Association Test (adapted from a similar diagram in Greenwald et al., 1998).

Figure A3: Graph showing the correlation between average score on the 'threat' sub-scale of the SAM against participants’ SA-IAT scores.

Figure A4: Graph showing the correlation between average score on the 'challenge' sub-scale of the SAM against participants’ SA-IAT scores.

Figure A5: Graph showing the correlation between average score on the adjusted 'threat' sub-scale of the SAM against participants’ SA-IAT scores.
List of Tables

Table 1  Means and SDs for PrS & PoS data for the Control group.

Table 2  Means and SDs for PrS & PoS data for the Experimental group.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Anterior Cingulate Cortex</td>
</tr>
<tr>
<td>AER</td>
<td>Automatic Emotion Regulation</td>
</tr>
<tr>
<td>AL</td>
<td>Allostatic Load</td>
</tr>
<tr>
<td>BPS</td>
<td>British Psychological Society</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary Heart Disease</td>
</tr>
<tr>
<td>DET</td>
<td>Differential Emotions Theory</td>
</tr>
<tr>
<td>ESES</td>
<td>Examination Self-Efficacy Scale</td>
</tr>
<tr>
<td>GI</td>
<td>General Intelligence</td>
</tr>
<tr>
<td>GSES</td>
<td>General Self-Efficacy Scale</td>
</tr>
<tr>
<td>GSR</td>
<td>Galvanic Skin Response</td>
</tr>
<tr>
<td>HPA</td>
<td>Hypothalamus-Pituitary-Adrenal</td>
</tr>
<tr>
<td>HRV</td>
<td>Heart Rate Variability</td>
</tr>
<tr>
<td>IAT</td>
<td>Implicit Association Test</td>
</tr>
<tr>
<td>JCQ</td>
<td>Joint Council for Qualifications</td>
</tr>
<tr>
<td>Ofqual</td>
<td>Office of Qualifications and Examination Regulations</td>
</tr>
<tr>
<td>PFC</td>
<td>Prefrontal Cortex</td>
</tr>
<tr>
<td>PrS</td>
<td>Pre-Stressor Testing</td>
</tr>
<tr>
<td>PoS</td>
<td>Post-Stressor Testing</td>
</tr>
<tr>
<td>PSS10</td>
<td>10-item Perceived Stress Scale</td>
</tr>
<tr>
<td>PTG</td>
<td>Post-Traumatic Growth</td>
</tr>
<tr>
<td>SA-IAT</td>
<td>Stressor Appraisal Implicit Association Test</td>
</tr>
<tr>
<td>SAM</td>
<td>Sympathetic-Adrenal-Medullary</td>
</tr>
<tr>
<td>SE</td>
<td>Self-Efficacy</td>
</tr>
<tr>
<td>SELF</td>
<td>Simple Ego-Type Life Form</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual Reality</td>
</tr>
</tbody>
</table>
List of Appendices

Appendix A  The effect of SE on academic performance.

Appendix B  Developing, Testing and Validating the Stressor Appraisal Implicit Association Test (SA-IAT): an Institution Focused Study.

Appendix C  Permission to use the GSES.

Appendix D  The Examination Self-Efficacy Scale (ESES).

Appendix E  The 10-item Perceived Stress Scale (PSS10).

Appendix F  Confirmation letter from the Research Ethics Committee.

Appendix G  Participant Briefing

Appendix H  Participant Consent Form
Buffering Preconscious Stressor Appraisal: the Protective Role of Self-Efficacy

Chapter 1: Background

The vast majority of contemporary theories of stress stem from the pioneering work of Walter Cannon, whose ground-breaking research in the early years of the last century established that emotionally salient stimuli were capable of eliciting change in physiological systems. From these initial publications a substantial and diverse field of research has blossomed, which contributes many thousands of papers every year – enough to more than fill the twenty or so journals which cater specifically for the subject. Given the exponential interest that Cannon's initial investigations engendered, it would be tempting to assume that stress was a modern phenomenon, perhaps borne from an increasing culture of individualism and ever-increasing expectation (Eckersley & Dear, 2002), or from the mounting zeitgeist of materialism (Twenge et al., 2010). However, although references to stress do disappear from the academic record before the mid-1920s, this is simply a consequence of etymology¹: in fact, history documents ancient incidences of stress: “in the Talmud and the Bible, we read that families have been affected by the events of change, trouble, disaster and ambiguity since the beginning of recorded time” (Boss, 1987, cited in Weber, 2011, pg696). Stress, therefore, is anything but new, despite its relatively recent emergence onto the academic stage.

¹ Stemming originally from the Latin word stringere - to draw tight - Walter Cannon first applied the word 'stress' to a physiological context in 1926.
Whilst it is tempting to dismiss such claims as alarmist, or a product of a deepening understanding of an expanding field of research; nevertheless, a growing number of authors now assert that stress is, rather like obesity, very much a growing problem internationally, particularly for young people (Zimmer-Gembeck & Skinner, 2008).

"In recent years there have been major increases in stress-related disorders in young people, including suicides, substances abuse, depression, anxiety and eating disorders... Such patterns have emerged world-wide and represent a challenge to policy-makers, service providers and families alike" (McNamara, 2000, pg xv - introduction).

The evidence substantiating such statements is significant: for example, comparisons of self-reports taken in 1950 and 1990 reveal a rise in anxiety levels of approximately one standard deviation. In fact, “anxiety is so high now that normal samples of children from the 1980s outscore psychiatric populations from the 1950s” (Twenge, 2000, pg1018).

Similarly, young people score about a standard deviation higher on clinical scales of psychopathology than their grandparents did at the same age, which includes a five-fold increase in the number of adolescents scoring above the thresholds for psychopathological diagnosis (Twenge et al., 2010). In a similar way, the PSS10 instrument I introduce later in this study displays a clear two-point increase in normative stress scores between 1983 and 2006, with further rises recorded in the last three years (Cohen & Janicki-Deverts, 2012). Although, like Wainwright and Calnan (2000), I resist the metaphor of a 'stress epidemic' (Claxon, 2008), it is indubitable that stress has become a national concern, particularly for young people, for whom approximately one quarter will
experience a significant life stressor during their formative adolescent years (Zimmer-Gembeck & Skinner, 2008) and as many as one student in two will experience symptoms of stress whilst they are at school (Chamberlain et al, 2011). It is against this disquieting backdrop that the rationale for this study is contextualised: if my ultimate aspiration for stress research is to facilitate the ‘inoculation’ (Howard et al., 1975) of adolescents against this ‘epidemic’, it is critical first to establish a comprehensive understanding of the pathogenesis of stress, so that programmes of intervention can be both informed and strategically structured to avert stress in young people before it becomes chronic.

At this point personal experience develops some relevance: having taught in British independent schools for the last decade, I have not only witnessed first-hand the harmful effects of stress, but also observed (happily and intriguingly) a number of students who, despite considerable academic pressure and/or unfortunate personal circumstances, did not develop stress. This baffling observation served not only as an emotive catalyst in prompting me to try to understand the sequence of events, or phenomena, that appeared to have protected those students from stress, but it also encapsulated the inceptive experience that inspired this study.

Undoubtedly many factors lie behind the genesis of stress - far too many to review here. However, amongst the numerous schema, models, hypotheses and suggestions, I have come to believe that efficacy theories offer the most potential to develop the next steps in stress research; both in terms of explaining the ‘anti-stress’ effect I described earlier, but also for the advancement of pedagogy, policy and practice that could successfully reduce the intensifying occurrence, and the
pathological consequences, of stress in young people. Simply put, our efficacy beliefs, or sense of self-efficacy, are the “measure of confidence or conviction that one can successfully execute the behaviour required to produce an outcome” (Bandura, 1977, pg79). I withhold advancing here the evidence emphasizing the important role self-efficacy (SE) plays in protecting adolescents from stressors – that comes later – however, I will point out that the popular media is replete with examples of the positive effects of SE in high stress scenarios. Athletes, for example, often ascribe their success to their belief in themselves - “some people say that I have an attitude; maybe I do. But I think that you have to: you have to believe in yourself when no one else does, that makes you a winner right there” (Venus Williams, cited in ITF, 2013, pg3). Similarly, after successfully scaling Everest in 1953, Edmund Hillary famously quipped that “it is not the mountain we conquer, but ourselves” (Hillary, cited in DeVyre, 2002, pg1), alluding both to the mental challenge of climbing the world’s tallest mountain, but also to the most significant resource that enabled him to achieve this feat – his belief in his ability to do it. In other words, Hillary clearly understood that, whilst SE “does not necessarily ensure success..., self-disbelief assuredly spawns failure” (Bandura, 1997, pg77). Indeed, the seminal work of Albert Bandura, in collating over fifty years of research at the end of the last century, (virtually single-handedly) succeeded in replicating experimentally exactly this scenario – whilst people with high SE appear to be protected from the effects of stressors, people with low SE are anything but. His research led assuredly to the conclusion that:
"Students' beliefs about their efficacy to manage academic task demands... influence them emotionally by decreasing their stress, anxiety, and depression" (Zimmerman, 2000 - original emphasis, pg82).

Bandura's assertion about the protective nature of SE in overcoming stress is, like stress itself, not really anything new. For example, in Book V of the Aeneid, Virgil chronicles the triumph of Cloanthus - one of Aeneas's captains - in a boat race to mark the first anniversary of Anchises's death. Virgil describes Cloanthus's victory with the phrase 'possunt, quia posse videntur', which translated literally means "they can, because they think they can" (Eden, 1993, pg168). Whilst this is, effectively, the same point that Bandura was making 2000 years later; what separates Bandura's eminent work from Virgil's social observation is; a) the weight of evidence Bandura collated, which more than fills the six books he published on the subject, b) his use of post-positivist epistemology to demonstrate a causal link between SE and stress and, c) his commitment towards the development of a model of social cognition, which could both explain, and then predict, the effect. In essence these degrees of separation form the grounding principles that shape the foundation of this research. Following Bandura, I too have chosen to explore the stress-reducing effect of SE within the post-positivist ontology: not only does this make available methodologies that seek to demonstrate causality between experimental variables (Shadish et al., 2002), but also, through the ontological axiom of objective, critical realism (Guba & Lincoln, 1994), such methodologies seek to accumulate knowledge of the world through the principle of generality. As I stated earlier, one of the guiding motivations behind this work is a desire to contribute, albeit in a very small way, towards the development of a
national curriculum that incorporates a focus on stress resistance, or anti-stress training. Without external validity, the conclusions of this study cannot be applied to such aims; which, simply in itself, warrants the choice of a post-positivist approach. A further consideration of following a Bandurian perspective is that a great deal of his formative research was conducted within an educational context. Given this, and the implicit objectives of the study, I made two key decisions early in the inception of this project:

1) To pursue an EdD (as opposed to a PhD), and thus to develop a significantly greater focus on the context of my research.

2) To prioritise the ecological validity of my research.

A necessary consequence of these decisions was that the methodological rigor of the study was, to a degree, impinged in some areas (most particularly when forced to choose between an experimental or a quasi-experimental methodology). However, in keeping with Bandura’s earlier work, the outweighing benefit of preserving ecological validity was that the study maintained a realistic, natural research milieu which, I argue later, lends considerably to the generality of my conclusions and, therefore, to my over-arching ambition of completing a rigorous research project that adds significantly to the pursuit of reducing stress in adolescents.

Where the study differs from Bandura’s prior work is part c) - the development of a model. Commitment notwithstanding, a fundamental vergence between this study and that of previous research is that I have
focused entirely on the events that occur in the first few seconds of exposure to an academic stressor. Effectively, this precludes a coping-based approach as, although I have no doubt that coping has a very significant role to play in the genesis of stress (e.g. Folkman, 1984; Dweck, 1999), I believe that recent breakthroughs in the study of psychopathologies such as addiction, anxiety and depression, have considerable scope to inform our understanding of stress. Therefore, I adopt a dual process perspective in this study (e.g. Compas, 2004) and examine the extent to which SE appears to protect adolescents from stress by exerting an effect on stressor appraisals, particularly those appraisals that occur preconsciously (Bargh, 1990; Bargh et al. 1993). To disclose any further details here would detract from the bearing of the study; however, as a final comment I wish to add that, if anything, the work that went into completing this project has served thoroughly to reinforce my conviction that, "among the types of thoughts that affect action, none is more central or pervasive than people's efficacy judgments" (Bandura, 1997, pg21), particularly in our response to stressors.
Chapter 2: Stress

For adults, a large body of research documents the serious negative impact of chronic stress on health, including the development of cancers (Adler, 1994; Antonova et al., 2011), predisposition to cardiovascular diseases (Kivimaki et al., 2006) including myocardial infarction (Rosengren et al., 1991), coronary atherosclerosis (Hollis et al., 1990) and stroke (Harmsen et al., 1990), depression (Kessler, 1997; Kopp et al., 2007), anxiety disorders (Stanford et al., 1993), infectious disease (Glaser et al., 1987; Stone et al., 1987), pregnancy complications (Pagel et al., 1990) and a number of other serious illnesses (Herbert & Cohen, 1993). In adolescents, the "exorbitant toll exacted from students by stressors" (Matheny et al., 1993, pg109) is also well documented and chronic stress has been strongly linked with poor academic performance (Barker, 1987), depression (McLaughlin & Hatzenbuehler, 2009) and may well contribute to violence and truancy (Fremont, 1993).

More worryingly, however, are the results from nation-wide studies, which document not only a rising incidence of stress-related self-harming in young people (Truth Hurts, 2005 - which showed that at least one child in fifteen self-harms, giving the UK the highest incidence of adolescent self-harming in Europe), but also that a growing number of school-aged children (approximately a million in the UK last year - National Centre for Health and Statistics, 2012) seriously consider taking their lives because of stress. This figure directly correlates with the rising incidence of teenage suicide (National Centre for Health and Statistics, 2012), which is now the third largest cause of death in young people (National Centre for Health and Statistics, 2012). Again, these studies have implicated a
causal role for stress, as high stress scores in self-assessment questionnaires are predictive of teenage suicide (Schneider, 2004; Green, 1984), particularly for adolescents exposed to academic sources of stress (Ang & Huan, 2006a; Nelson & Crawford, 1990;). Indeed, the recent observation that, globally, teenage suicide rates show a good degree of correlation with the timing of public examinations (Toero et al., 2001) - especially in cultures where examinations are perceived to be particularly stressful (Sue & Okazaki, 1990) - has led campaigners for young people’s health to denounce stress as the “children’s epidemic of the twenty-first century” (Claxon, 2008, pg192). Nationally, this voice is growing stronger as more evidence accumulates to substantiate the link between ill health, psychopathology and the ‘mounting anxiety’ and ‘intolerable academic pressure’ (Hough, 2011) increasingly suffered by adolescents at school (Twenge, 2000).

Unfortunately, however, the association between stress and mortality is not simply limited to suicide, as many studies have revealed other potential indirect links between stress and morbidity. For example, stress generated from natural disasters has been linked not only in an increase in the incidence of suicide (Krug et al., 1998), but also an elevated frequency of mental illnesses (Rubonis & Bickman, 1991), including depression (Ollendick & Hoffmann, 1982; Siegel & Brown, 1988), anxiety (Gjerde, 1995), insomnia (Maida et al., 1989) and symptomatology for a range of other ‘internalising’ pathologies (Compas et al., 1989) and affective disorders (Leadbeater et al., 1995). Additionally, exposure to stressors has also been found to correlate with various ‘externalising’ problems such as delinquency and aggression (Baer et al., 1987; Rutter, 1988).

---

2 Although see Bravo et al., 1990, who argues that these effects are pre-existing.
as well as an elevated rate of substance abuse and domestic violence (Shore et al, 1986). However, compounding these effects is the wealth of data supporting a direct and causal relationship between stress and early death. For example, Leor and colleagues observed an increase in the number of cases of sudden cardiac death in response to an earthquake (Leor et al, 2003), whilst Li’s research group observed an increase in mortality rate in family members who had suffered the death of a child (Li et al, 2003). Similarly, Eitinger (1973) found a strongly increased mortality rate in Norwegian concentration camp survivors, which he attributed to the effects of chronic stress exposure. Whilst it is easy to criticise such studies on the grounds that they rely on retrospective analyses, or employ small sample sizes, recent large-scale national and international prospective studies have conclusively established that stress does indeed decrease life span, particularly for males (Nielsen et al, 2008). This effect is likely to occur either through decreased immune functioning (Segerstrom & Miller, 2004; Kiecolt-Glaser et al, 1991) resulting in increased healing time (Kiecolt-Glaser et al, 1995) and greater susceptibility to disease (Kiecolt-Glaser et al, 1996), or through specific stress-induced pathologies, amongst which cardiovascular disease (Ohlin et al, 2004; Iso et al, 2002; Eaker et al, 2004), stroke (Rosengren et al, 2004), cancer (Ferraro & Nuriddin, 2006) and accelerated senescence (Epel et al, 2004) are specifically implicated. All of these conditions are aggravated by, but not limited to, stress-

---

3 These ‘externalising’ effects are significantly more pronounced in boys. By contrast, girls exhibit greater ‘internalising’ conditions (Compas et al, 2001).

4 It should be noted that this study has been criticised on the grounds of statistical validity as eight of the twelve analyses of mortality described in this study were not statistically significant (Rogers & Reich, 1988). Other studies have also failed to observe the same effect (e.g. Levav et al, 1988).

5 It is also worth noting that some studies with similarly small sample sizes have not found a link between stress and morbidity (e.g. Singer et al, 1976; Rasul et al, 2004).
induced changes in lifestyle, including the reduction in uptake of preventative behaviours (Schulz et al, 1997), increased smoking (Cohen et al, 1993), missing sleep and elevated consumption of alcohol (Cohen & Williamson, 1988).

Collectively, the implications of these studies are significant: although psychopathologies frequently reveal an augmented prevalence in adolescence naturally (Cohen et al, 1993; Dekovic et al, 2004 - and particularly for affective disorders Steinhausen et al, 1998), stress is known to amplify the effect markedly and, consequentially, many of these conditions also show strong co-morbidity with stressor exposure (Seiffge-Krenke, 2000). The supposition, therefore, is clear: "stress... constitutes a significant and pervasive risk factor for psychopathology in childhood and in adolescence" (Compas et al, 2001, pg87).

This disquieting conclusion is exacerbated by the observation that young people frequently experience stress at school (Compas, 1987; Currie et al, 2004) which, in the UK, may affect as many as one student in two (Chamberlain et al, 2011). Other surveys report even higher incidences of adolescent stress: Lohaus (1990), for example, interviewed 342 pupils in a range of secondary schools and found that 72% of the children interviewed had experienced significant academic stress within the previous month (either as situations related to school work, or stress related to other performance-oriented contexts) - a statistic that was further increased through additional stress experienced as a consequence of social problems, rather than from exposure to academic stressors. Reports indicate that, although many adolescents experience specific stressful events at school (Ng et al, 2003), repeated exposure to
minor stressors is a more significant predictor of psychopathological state (Siegel & Brown, 1988), rather than the occurrence of significant negative life events (Siegel & Brown, 1988; Seiffge-Krenke, 2000). These 'minor stressors' cover a considerable range of events and include both academic (Kohn & Frazer, 1986) and social (Sgan-Cohen & Lowental, 1988) stimuli. However, most studies conclude that, whilst social stressors have a significant effect on the health of adolescents, the pressure to achieve academically (Hirsch & Ellis, 1996) and, particularly, the significance of strong performance in examinations (Lakshmi, 2009) constitute the most salient stressors experienced by adolescents (Burnett & Fanshawe, 1997; Shah et al, 2009). Clearly, then, understanding (and then preventing) the mechanisms by which academic stressors engender stress in adolescents is of great importance, not only for the health and well-being of young people, but also to avert what the Chairman of the American Psychological Society has recently described as “our next public health crisis” (Anderson, 2010, pg17).

What is stress?

Stress is, in fact, pretty difficult to define and, consequentially, a large and varied range of descriptions and characterisations of stress exist in the literature. Butler (1993) deals with this problem by advocating thinking about stress in three different ways. She argues that stress can be encapsulated in terms of stimulus overload and draws the example of the collapse of a load-bearing beam. Stress, therefore, is a condition elicited by the combined 'pressure' of stressors - stimuli or events, which elicit the stress. Whilst it is important to make clear the distinction between stress and stressors, the notion of stress as a unidimensional
state - i.e. being 'stressed' or 'not stressed / normal' (Duffy, 1962) - is very unsatisfactory, chiefly because it omits any consideration of the uniquely subjective nature of stress (Lazarus, 1999; McEwen, 2007). Therefore, I move onto Butler's second model of stress, which views stress as a variable, adaptive response to stressors.

Based on the seminal work of Hans Selye, response-based theories tend to view stress in terms of "the nonspecific response of the body to any demand made upon it" (Selye, 1974, pg32). Current theories focus on endocrine responses to stressors\(^6\), particularly activation of the HPA and SAM axes, resulting in increased glucocorticoid and catecholamine secretion (see Schlotz [2008] for a review of these endocrine responses). Within this definition of stress, such hormones (and most particularly cortisol) mediate adaptive metabolic responses such as gluconeogenesis and glycogenolysis, which are collectively categorised as the acute response to stressors (McEwen, 2004). These acute allostatic responses are purported to protect against the impact of the stressor (although, under some circumstances, they are also acknowledged to have damaging effects on the body too - McEwen, 2005). By contrast, acute stress can be differentiated temporally and physiologically from chronic stress, which is defined in terms of maladaptive allostatic loading (McEwen, 2004 - see later), and tends to occur if exposure to the stressor is prolonged.

In support of response-based theories, perceived stress is reliably correlated with elevated cortisol levels (Dickerson & Kemeny, 2004;  

\(^6\) Traditionally, one of the most significant limitations of response-based theories of stress as they tend not to recognise, or under-represent, psychological or cognitive influences on stress.
although see Kirschbaum et al., 1999 who argue that it is not) and many stress-related diseases (e.g. Burke et al., 2005) are associated with elevated or abnormal glucocorticoid secretion. Thus, in response-based terms, stress may be defined as a physiological reaction to a stressor, and is now commonly viewed in terms of allostatic load (McEwen, 1998). Allostasis refers to the physiological mechanisms that maintain homeostasis\(^7\) (Romero et al., 2009) and allostatic load (AL) is the physiological cost of allostasis - i.e. the “cumulative result of an allostatic state” (McEwen & Wingfield, 2003, pg4) or:

“The wear and tear on the body and brain resulting from chronic over-activity or inactivity of [allostatic] physiological systems that are normally involved in adaptation to environmental challenge” (McEwen, 1998, pg37).

McEwen and Wingfield (2003) recognise two different groups of AL; Type I, which occurs when the energy required for allostasis exceeds that available from the environment, and Type II, which occurs when sustained activity of allostatic mechanisms produces maladaptive or pathological side-effects. McEwen (2000; 2007; see also Goldstein & Kopin, 2007) presents Type II AL as a model for stress, arguing that prolonged, inadequate, unbalanced\(^8\) or repetitive activation of allostatic systems causes chronic stress and stress-related illness.

---

\(^7\) Homeostasis refers to the physiological parameters (such as blood glucose concentration, or the degree of haemoglobin \(O_2\)-saturation) that must be held constant in order to sustain life (McEwen, 2000).

\(^8\) McEwen (2000) proposes that an inadequate response in one system leads to compensatory hyperactivity in other allostatic systems, which is what causes the damage.
McEwen’s proposal for AL as a mechanism for stress is well supported by experimental evidence, the most significant of which is the McArthur Study on Successful Ageing – a longitudinal study of 1189 high-functioning, healthy women aged 70 - 79. Starting in 1997, participants were interviewed each year and, using data from psychometric assessments, anthropometric protocols, blood and urine samples, AL was calculated quantitatively using data from a combination of primary endocrine mediators and secondary outcomes (e.g. blood pressure, hip-to-waist ratio etc). The study found that AL strongly predicted mortality (Seeman et al., 2001), was closely negatively correlated with cognitive and physical functioning (Steward, 2006) and exhibited independent explanatory power over and above a measure of doctor-diagnosed disease (Seeman et al., 2004), thus collectively implying a causal link between AL, mortality and disease. Similar conclusions were drawn by Goldman and colleagues (Goldman et al., 2006), who reported close correlations between AL and poor mental and physical functioning in a prospective study of morbidity in middle-aged Taiwanese people. Crimmins has also observed the same effect in the USA across a wider age bracket (20 - 50), which remained statistically significant even after age and gender had been controlled for (Crimmins et al., 2006). Others reliably replicate the McArthur study, leading to the suggestion that the correlation between biomarkers of AL and mortality may be causal (Gruenewald et al., 2006; Karlamangla et al., 2002). Taken together, these studies provide persuasive evidence in support of response-based, physiological models of stress, most particularly as they have been found to be excellent predictors for the onset of stress-related disease.
That said, there are a number of limitations with response-based models. Firstly, they tend to focus heavily on glucocorticoid activity, ignoring the conclusions of researchers like Taylor (2006), whose work strongly implicates other hormones as mediators of stress. Secondly, they place a good deal of emphasis on energy budgets in the development of stress-induced pathology (Romero et al., 2009) and evidence linking cortisol with increased energy expenditure is conflicting (Remage-Healey & Romero, 2001; Cyr et al., 2008). Thirdly, they fail to differentiate between specific stressors, and thus view all threats to homeostasis as stressor stimuli of equal valence (Day, 2005), which is clearly not the case, for example, for a glucose load after a heavy meal and a close encounter with a runaway bus! Finally, as with condition-based models of stress, response-based models also leave restricted scope to account for the oft-reported subjectivity of stress, as it is well-known that people respond to stressors in very different ways (Lazarus, 1991; 1999), even when the stressors they experience are identical (e.g. see Williams et al., 2011). Therefore, although response-based models such as AL have contributed significantly to our understanding of the genesis of stress-related disease, I posit that they do not offer sufficient explanatory potential to account for the individualistic nature of stress. Therefore, I move onto Butler's (1993) third and final group of definitions, which view stress as a dynamic relational process (Lazarus & Folkman, 1984).

According to this view, “stress is more than a physiological response to environmental demands: it is based on personal perception and individuality” (Lazarus & Folkman, 1984, pg19), which may explain both the individualistic experience of stress (Cooper & Bright, 2001) and the wide

---

variation in responses that occur when different people are exposed to the same stressor (Williams et al., 2011). Research in this field began through studies with war veterans, whose widely varying experiences of stress in combat (Lazarus & Erikson, 1952) led Richard Lazarus to the joint conclusions that, 1) stress is experienced very subjectively and, therefore, 2) that stress originates in the cognitive interpretation of stressors - a process he referred to as appraisal (Lazarus, 2001; 1999; 1993 see also Lazarus & Lazarus, 1994; Lazarus & Folkman, 1984). Developed in the late 1960s, Lazarus's Cognitive Mediation theory (Lazarus, 1966; 1968) evolved to view stress as a subordinate component of emotion (Lazarus 1993; 1999; 2001 see also Smith & Kirby, 2011) - "psychological stress should be considered part of a larger topic, the emotions" (Lazarus, 1993a, pg10). In this view, stress exists within a "part-whole relationship" with affect (Lazarus, 1999, pg37), leading to its definition in functional terms as an affective process, or an emotion. Thus, so the argument goes, as an emotion stress serves the adaptive function of "mobilis[ing] the organism to deal quickly with important interpersonal encounters" (Ekman, 1992, pg171) and, therefore, acts as:

"The interface between an organism and its environment mediating between constantly changing situations and events and the individual's behavioural responses" (Scherer, 1982, pg556).

In support of considering stress in affective terms, I suggest three unique advantages of this position over the other approaches outlined earlier: 1) it may begin to explain the subjective nature of stress which, as I asserted above, both condition- and response-based definitions do not adequately account for; 2) it reinforces (and may even predict) the
success cognitive intervention programs have achieved in limiting the stress experienced by adolescents at school (Hampel, Meier & Kummel, 2008; Mori & Uchida, 2009; Frydenberg, 2004; Kraag et al., 2006; Lohaus, 2011); and 3) it promotes the conception of stress from a functional standpoint, rather than a physiological end-point, which is arguably more useful from a pragmatic, pedagogical or clinical perspective. Apropos, in functional terms, emotions are:

"Unique experientially-valenced 'state spaces' that help organisms make cognitive choices - e.g., to find food when hungry, water when thirsty, warmth when cold, and companionship when lonely or lusty...: they [emotions] provide intrinsic values - organic 'pressures' and 'drives' for the guidance of behaviour" (Panksepp, 2003, pg6).

Therefore, to view stress as an emotion (as I advocate) one should expect to see evidence of adaptive stress-induced changes in cognition and behaviour. Fortunately, such evidence is plentiful! However, before substantiating my assertion that stress is an emotion, I first wish to outline briefly what I mean by 'emotion'.

**What are emotions?**

Defining stress as an emotion leads immediately to the problem of establishing a definition for emotion. This is no easy task as a plethora of different descriptions and characterisations exist in the literature: for

---

10 Adopting an emotional definition of stress also allows considerably more scope for the development of further intervention programs such as those cited here, thus shifting the emphasis in stress-prevention away from drug treatment (which tends to be the outcome of physiological response-based models like AL) and towards vastly more practical (and, arguably, more successful) pedagogic interventions.
example, Kleinginna and Kleinginna (1981) identify nearly one hundred different definitions of emotion and that number has certainly risen over the three decades that have passed since they published their study. This issue is further complicated by the use of ambiguous terms such as 'mood', 'disposition' or 'feeling' (Scherer, 2000) which are employed very differently by different authors, often to mean a temporary predisposition to emotion, or to a low intensity emotion (Oatley & Johnson-Laird, 1987), but even this is inconsistent across the literature!

However, despite these difficulties, I have managed to draw together commonalities from the field that have helped me to bring together a working definition (Frijda, 1986) of 'emotion.' Therefore, I posit that emotions;

1. Are manifested within the brain (Davidson et al., 2009; LeDoux & Phelps, 2008; LeDoux, 1998; Rolls, 1999; Lane et al., 1998; Panksepp, 1991; 1998), either via emergent constructivist processes (Barret, 2009; Lindquist et al., 2012) or through the activity of bespoke anatomical structures (Panksepp & Watt, 2011; Ekman & Cordaro, 2011) or networks (Izard, 2011).

2. Are neurocognitive states (Ekman, 1992; Frijda et al, 1991) or processes (Scherer, 1982) which alter behaviour and cognition.

---

11 See Reisenzein's (2007) commentary on the notion of nominal vs real definitions in research.

12 Bottom-up somatovisceral feedback undoubtedly has a pronounced effect on emotion (Cacioppo et al., 2000; Larsen et al., 2008). However, I do not accept the James-Lange view of independent visceral emotions; instead I posit that somatovisceral information biases or sets the tone of higher level substrates (Berntson, Sarter & Cacioppo, 2003) and thus does not generate emotion per se.

13 By defining emotions as having neurobiological origins, I dismiss social constructionist theories of emotion (e.g. Mesquita's [2010] theories of social context or Harré's [1986] and Averill's [1980] constructionist model).
adaptively (Pinker, 1997; Panksepp, 2003) to optimise intrapersonal functioning (Levenson, 1999; Keltner & Gross, 1999)

3. Are not single modalities, but rather consist of multiple 'coherent clusters of components' (Silva, 2006), which include physiological regulation, motor expression and subjective feeling\(^{14}\) (Scherer, 1993) and may also include motivation and cognition (Moors, 2009; Scherer, 2000).

4. Are episodic (Scherer, 2000) and can be characterised by an abrupt onset, which fades away over time (Frijda et al., 1991; Ekman, 1992).

5. Are triggered by internal or external stimuli or events that are of major significance to the organism (Frijda, 2008). These stimuli are identified as such through appraisal processes, which include cognitive appraisal mechanisms (Lazarus, 1991; 1999; Frijda, 1986; 1987; Roseman, 1991; Smith and Ellsworth, 1987) and preconscious processes (Zajonc, 2000; 2001; LeDoux, 1998, LeDoux & Phelps, 2008) in a multi-level appraisal system (Dahaene et al., 2006).

Beyond these five points of agreement a very diverse range of definitions and models of emotion exists - certainly far too many to review here. My view is that emotions are best considered from a multi-level perspective (Robinson, 1998), in which;

---

\(^{14}\) The reaction triad of emotion (Scherer, 2000).
"Emotion is the product of an interaction between simple, non-conscious, automatic processes and deliberate, conscious and controlled processes" (Ochsner & Barrett, 2001, pg39).

This view depicts emotions as either a two- (Levenson, 1999) or three-tier (Panksepp & Watt, 2011) hierarchical neurobiological system, with an underlying and largely automatic\textsuperscript{15} core processor of basic emotions (Panksepp 1998; Ekman & Cordaro, 2011) which initiate (or part-comprise) adaptive states of 'action readiness' (Frijda & Parrott, 2011), or 'modes of adaptation' allowing people to adapt efficiently to changing environmental demands (Levenson, 1994). Panksepp (1998; 2001) identifies seven such universal (Ekman, 1992) and evolutionarily-primitive (Panksepp, 2005) basic emotions in the core system (seeking, rage, fear, lust, care, panic and play), which he views as subcortical and, therefore, largely automatic. Evidence from neuroimaging meta-analyses largely supports the existence of these basic emotional systems (Denton, 2006; Vytal & Hamann, 2010; see also Murphy et al., 2003; Phan et al. 2002), which is further corroborated by Mobbs and his colleagues (Mobbs et al., 2007), who observed a clear graded shift in emotional brain activity from 'higher emotion' in cortical centres to 'basic emotion' in subcortical areas in response to increasing threat (a 'virtual predator' in a VR maze). This is comparable with evidence from animal brain imaging studies (Logothetis et al., 1999) and positive and negative reinforcement animal learning paradigms (Dawkins, 2006) which, in conjunction with other research on animal emotions (Griffin & Speck, 2004; Beshkar, 2008), provides

\textsuperscript{15} I define automatic processes as those that are conducted unconsciously and without direct cognitive attention or intention (Bargh, 1994; see also Koole & Rothermun, 2011). It should be noted that this definition does not mean that automatic processes are hard-wired or inflexible, quite the opposite, in fact (see later section on preconscious appraisals)!
persuasive evidence to support the existence of the evolutionarily-conserved basic emotion system Panksepp (1998) advocates\textsuperscript{16}. Although these basic programmes are largely conceived to be automatic (Levenson, 2011), it has been suggested (e.g. Ekman, 1984) that input from higher cortical systems adaptively modulates these core programmes and, when combined with evidence of subconscious associative and operant learning within the core emotions (e.g. LeDoux’s [1998] work on fear conditioning in the amygdala or Öhman’s [1993] studies on anxiety in humans\textsuperscript{17}), the conclusion I propose is that significant malleability occurs within these ‘core’ neurobiological affective systems.

In tandem with these core emotions, I support the existence of:

“A more recently evolved, highly flexible, and much less predictable set of control mechanisms that are designed to influence the actions of the core system” (Levenson, 1999, pg483).

Numerous evidence from animal electrophysiology experiments (Rolls, 1999) and neuroimaging studies (Damasio, 1994) implicates specific neocortical areas\textsuperscript{18} in the regulation of higher emotion - a finding reliably corroborated by patients with cortical lesions, who have been found to display emotional deficits but, crucially, do not appear to suffer complete loss of affective reaction (Damasio, 1994; see also Berridge, 2009). Thus, I advance that a very wide range of cognitive processes may contribute towards emotion by channelling their collective effect into the regulation

\textsuperscript{16} It is also worth highlighting that evidence also exists supporting the evolutionary phylogeny of these systems (see, for example, Cabanac, 1999).

\textsuperscript{17} See also Panksepp (2001) for associative learning in other core command systems.

\textsuperscript{18} Particularly the left and right PFC and ACC.
of a finite number of core subcortical emotional systems, which are thought to generate the underlying affective state (Panksepp, 1998). According to this view, higher cortical systems, therefore, regulate the underlying core emotions (Davidson et al., 2000; Berkman & Lieberman, 2009) as well as generating affective states directly. This regulation is believed to be important, not only because it may elicit significantly greater flexibility in the control of affect, but also because - as I explain later in this Chapter - it is one of the significant points of congruence with stress research. Before looking at the (sizeable) cross-over that exists between stress research and theories of emotion, I wish briefly to outline what I mean by emotion regulation before looking at the evidence that supports my assertion that stress is best considered in functional terms as an affective process.

**Emotion Regulation**

Emotion regulation “involves the initiation of new, or the alteration of ongoing, emotional responses through the action of regulatory processes” (Ochsner & Gross, 2005, pg242). These processes can be roughly divided into two separate groups: conscious emotion regulation and automatic emotion regulation (AER).

**Conscious emotion regulation** processes are often separated into antecedent-focused strategies, which, “occur before appraisals give rise to full-blown emotional response tendencies” (Gross & Thompson, 2006, pg12) and response-focused strategies, which are defined as occurring after the responses are generated (Gross, 1999; 2001). Antecedent-focused strategies are known to be capable of decreasing the experience
of negative emotion with little physiological arousal or impact on memory (Ochsner & Gross, 2005) and have been further characterised into four sub-groups; situation selection, situation modification, attentional deployment and cognitive change. Of these, the first two strategies are usually couched in terms of the regulation of emotion via deliberately absenting (or exposing) oneself to places and situations with known associations with specific emotive stimuli, thereby deliberately activating (or seeking to avoid) specific affective states. In a similar way attentional deployment is defined as purposely distracting oneself from (Stifter & Moyer, 1991) or explicitly concentrating on (Wegner & Bargh, 1997) a specific stimulus - in effect “an internal version of situation selection” (Gross & Thompson, 2006, pg16). Cognitive change is subtly different, it has been described as:

"Changing how one appraises the situation one is in so as to alter its emotional significance, either by changing how one thinks about the situation, or [changing how one thinks] about one’s capacity to manage the demands it poses" (Ochsner & Gross, 2008, pg154).

To date, three categories of cognitive change have been identified;

1) Stimulus reappraisal (Ochsner, 2007), in which a stimulus is cognitively transformed so as to alter its emotional impact (Gross, 1998a). Evidence suggests that this occurs via a top-down route to ‘tone up’ or ‘tone down’ the emotional significance of a stimulus in order to re-represent the meaning of that stimulus in a goal-congruent way (Ochsner, 2007).
2) Cognitive repression of affect (Gross, 2002), which is purported to play an adaptive role by deliberately reducing a negative emotion (Ochsner, 2007). However, this potentially important aspect of response-focused emotional regulation may also come at some psychological cost (Goldin et al., 2008) and has been implicated with maladaptive psychological and physical functioning in the long term (Abelson et al., 2005).

3) The controlled generation of emotion, including the anticipatory elicitation of emotion in expectation of pain (Hsieh et al., 1999; Ploghaus et al., 1999) and in scenarios when memory leads to the negative appraisal of a normally neutral stimulus (Ochsner & Gross, 2005). Both examples reinforce the view that, not only does the cognitive interpretation of the stimulus appear to be critical in determining affect, but also (particularly in the case of Ochsner and Gross's [2005] study) that the effect is most likely mediated through top-down inhibition, or activation of subcortical systems. Both of these scenarios support a model of emotion in which affect arises from the interaction of top-down regulatory processes (including appraisals and reappraisals) on underlying subcortical systems (Ochsner & Barrett, 2001; Ochsner & Gross, 2005; 2008; Ochsner, 2007; Barrett et al., 2007; Gross & Thompson, 2006; Gross & Barrett, 2011).

Automatic emotion regulation systems appear to exhibit significant plasticity (e.g. Öhman & Soares, 1998) and much evidence supports their modulation by top-down conscious processes (e.g. Kunde et al., 2003; Kentridge et al., 2004; Dahaene et al., 2006) - i.e. as described in the previous section, conscious processing and/or prior experience seem capable of altering the way emotive stimuli are appraised preconsciously.
However, it has recently been suggested that modulation may also run in the opposite direction too. In other words, preconscious systems could well influence the functioning of neocortical emotional centres (Ferguson et al., 2005). Over the last decade research in this area has expanded dramatically and the field of implicit, or automatic emotion regulation (AER - see Mauss et al., 2007; Gyurek et al., 2011) is the result.

AER can be defined as processes:

“That operate without the need for conscious supervision or explicit intentions, and which are aimed at modifying the quality, intensity, or duration of an emotional response. Implicit emotion regulation can thus be instigated even when people do not realise that they are engaging in any form of emotion regulation and when people have no conscious intention of regulating their emotions” (Koole & Rothermund, 2011, pg390).

A wide body of research has implicated processes which automatically regulate attention (such as attentional tunnelling [for a review see Staal, 2004] or the positivity effect [Isaacowitz et al., 2006]) as playing a significant role in AER. As such processes “resolve emotions before they are fully triggered” (Mauss et al., 2007 - original emphasis, pg149), they are often referred to as antecedent-focused AER, with a similar semantic (Gross, 2001) and functional (Mauss et al., 2007) emphasis to conscious regulatory strategies. Although response-focused AER systems have also been identified (i.e. the regulation of behaviour and affective state “after the emotional response has been triggered” [Mauss et al., 2007, pg154], which include cognitive engagement, cognitive disengagement and behavioural regulation), the majority of research into
AER has tended to focus primarily on antecedent-strategies because of their potential to promote adaptive behaviour and cognition.

For example, many researchers (e.g. Custers & Aarts, 2005; Bargh & Gollwitzer, 1994; see also Bargh & Williams, 2007) have found that explicit performance goals can be activated and executed without conscious volition or awareness. Bargh, for example (Bargh et al., 2001), used an implicit priming methodology to prime participants subliminally with cooperative goals on a problem-solving task. When performance on the task was measured Bargh found that, not only did primed participants out-perform control groups, but that they did so by cooperating with each other. Crucially, the adoption of cooperative goals occurred without the participants’ explicit knowledge that they had acted in that way - i.e. they were unable to explain accurately why they had performed well. Bargh concluded that unconscious goal pursuit had occurred and, therefore, that, under some circumstances, preconscious systems may have access to mental representations of ‘desired outcomes’ - i.e. goals. Koole and Rothermond (2011) go further and suggest not only that preconscious processes have access to goal-related memory, but that they are capable of eliciting goal-targeted behaviour without conscious awareness. They propose a preconscious stimulus reappraisal process similar to cognitive reappraisal to explain the phenomenon of unconscious goal pursuit. In this view:

"People are not aware of any strategic attempts to create the reappraisal, [and thus]... people may easily perceive it as a true picture of the given situation. Reappraisals resulting from automatic processes and biases may thus have a much
higher chance to change personal beliefs and resulting emotions” (Koole & Rothermond, 2011, pg392).

A similar view is held by Silvia Bunge, who cites repressive coping as an example of automatic stimulus reappraisal (Mauss et al., 2007). Bunge asserts that automatically re-evaluating a negative emotive stimulus as being less stressful than it would normally be appraised, plays a significant adaptive role in helping people recover from exposure to serious life stressors, such as the death of a spouse (Bonanno, 2005). The implication being that the preconscious reappraisal processes that are believed to mediate repressive coping serve to protect the individual by diminishing (or, possibly extinguishing completely) the valence of the emotive stimulus, thereby reducing the degree of activation of the ensuing affective state. As I shall explain, this suggestion not only sketches out a very powerful role for cognitive resources (such as coping strategies) in preventing the development of chronic affective state (and thus, potentially, psychopathology), but it also draws tantalising links between the study of emotion regulation and stress research by placing AER and coping within the same functional bracket. As I asserted earlier, I believe that the connection between emotion and stress Bunge alludes to in repressive coping stems from their equivalence as affective processes and, in the following section, I set out the evidence that reinforces my position.

Evidence that stress is an emotion

1. Evidence from neuroimaging studies reliably identifies activity in identical brain areas during emotion regulation tasks and during episodes
of psychosocial stress (e.g. see Wang & Saudinio, 2011). For example, in a public speaking task Tillfors and colleagues (Tillfors et al., 2002) observed significant increases to cerebral blood flow in the prefrontal cortex (PFC) and anterior cingulate cortex (ACC) - neocortical areas specifically implicated in the regulation of emotion (Ochsner et al., 2002; Davidson, 2003). Similar patterns of activity were detected in other stress-related tasks, such as Stroop tests (Gianros et al., 2005) and combined stress tasks (Kern et al., 2008). By itself, this evidence is inconclusive as stress is commonly associated with negative emotional states (Shina, 2001) and, more recently, has been found to accompany positive emotions as well (Folkman, 2008). Therefore, whilst neuroimaging studies do support my assertion that stress is an emotion, the activity of brain emotion regulation systems during stress could be explained separately in terms of regulating stress-induced positive or negative emotions.

2. Evidence from studies on animals with cortical lesions clearly identifies a role for the neocortical areas that regulate emotion\(^{19}\) in activating the physiological stress response (Feldman & Conforti, 1975), possibly via a neurological ‘stress circuit’ (Cerquiera et al., 2008)\(^{20}\). Activation of this circuit appears to be highly stimulus-specific (Amat et al., 2005; Spencer et al., 2004), implying that the cognitive appraisal of stressors most likely makes use of similar neocortical circuitry that appraises other emotion-generating stimuli. Such a conclusion is further

\(^{19}\) In this case, specifically the medial PFC.

\(^{20}\) Interestingly, the stress circuit appears to be able to differentiate between intrinsic ‘systemic’ stressors (which activate the HPA axis automatically) and ‘processive’ stressors, which require appraisal by neocortical areas before HPA activation (Herman & Cullinan, 1997).
supported by neuroimaging studies, which implicate the PFC and ACC in cognitive appraisal of stressors (Johnson et al., 2002), even when selective attention to stimuli was controlled for (Kalish et al., 2006).

3. **Studies employing people with neocortical damage** further support an overlap between stress and emotion, as studies investigating the impact of neocortical damage report not only attenuated stress in patients with such lesions (Diorio et al., 1993; Buchanan et al., 2010) but also disturbed emotion regulation (Bechara et al., 1994; Anderson et al., 2006) as well.

4. **Studies highlighting the effects of stress on cognition** imply that, as an emotion, the stress process may be able to modulate cognition adaptively to optimise intrapersonal functioning. For example, it is reliably observed that experience of stress results in increased attention to stressors (Bradley et al., 1998; Compton et al., 2003) in a manner that selects attention towards specific stimuli known experientially to have high stress valence (MacLeod et al., 2002). This seems also to produce dramatic improvements in the memory of stressors (e.g. Bohannon's [1998] work on flashbulb memories after the Challenger shuttle disaster, or Cahill's [Cahill et al., 1994] demonstration that enhanced memory of emotional events is attenuated by drug-induced blockade of stress hormones), which mirrors effects on memory reported with other emotions (e.g. see Roozendaal et al., 2009). This memory-enhancing effect is posited to occur biphasically (i.e. an initial positive

---

21 The notion of stress as a positive functional process is not new as eustress was a central feature of Selye's (1974) original GAS theory. However, the concept of eustress seems to have fallen out of favour somewhat and has only recently become a focus of investigation.

22 A processes Compton and colleagues (Compton et al., 2003) refer to as hypervigilance.
effect, followed by attenuation and negative outcome if the stressor exposure is chronic) via the differing long-term effects of glucocorticoids on both glucocorticoid receptors and mineralocorticoid receptors in the brain (Ron de Kloet et al., 1999) - a theory which also accounts for why most research so far has tended to find only negative effects of stress on cognition (e.g. Kim & Yoon, 1998; Kirschbaum et al., 1996), particularly as animal (and particularly human) research methodologies tend not to examine the differing temporal effects of glucocorticoids in separate brain areas across two very different receptor systems.

Intriguingly, it may be that chronic or traumatic stress may also hold adaptive value. A recent body of research has begun to emerge documenting the phenomenon of post-traumatic growth (PTG - see Tedeschi & Calhoun, 2004), in which “an individual experiences positive psychological change (above pre-stress levels) as a result of the struggle with trauma or any extremely stressful event” (Jackson, 2007, pg2). This effect appears to manifest as an increased appreciation of life, enhanced sense of new possibilities in life, increased personal strength, improvement in close personal relationships and positive spiritual change (Tedeschi, Park & Calhoun, 1998). At the moment there is no clear consensus on the underlying cause of PTG (or, indeed, whether it is artefactual or a de facto phenomenon in its own right) but, if authentic, PTG provides compelling evidence for the adaptive role of stress and, therefore, very much supports my definition of stress as an emotion with adaptive function (Keltner & Gross, 1999).

23 Which are much less frequently distributed, but have approximately ten times more affinity for cortisol (Ron de Kloet, 1991).
5. The early development of the cortisol response in infancy also supports my view that stress is best viewed as an emotion. Although it is arguably inappropriate to use the term 'stress' (as perceived by an adult) as applied to infants, there is clear evidence that HPA regulation and cortisol response to stressors becomes established within the first year of a baby's life (Gunnar & Donzella, 2002) and that successful attachment with a primary caregiver plays an important part in this process (Gunnar et al., 1996). The establishment of HPA responsiveness and diurnal cortisol rhythm constitutes a crucial developmental stage through which normal regulation of the HPA axis is established (Graham et al., 1999) and the adaptive stress process becomes activated. This is essential, not only for enabling infants to experience stress (which, as I have argued earlier, exerts adaptive functionality), but it also proves critical in establishing normal affective regulation in later life (Hammen et al., 1992) and evidence is mounting to support a role for early HPA activity in avoiding pathological states such as depression in later life (Browne & Finkelhor, 1986). Therefore, for a normally-attached infant, early experience of controlled stress (e.g. brief separation from the primary caregiver) seems to serve a particularly important role in normal emotional and psychological development. The central tenet of the Differential Emotions Theory (DET – Izard, 1978; 1991; Izard et al., 1995) is that:

"Emotions retain their adaptive and motivational functions across the lifespan, [but] different sets of emotions become relatively more prominent in the different stages of life as they serve stage-related developmental processes" (Abe & Izard, 1999, pg523).
Thus, in regulating the development of the HPA axis (and, potentially, later mental health), the infant stress process plays out precisely the kind of stage-related developmental function posited by DET. Thus I assert that, in regard to its significance developmentally, stress is functionally homogenous to the other emotions in DET and, therefore, ought to be viewed with equivalence as an affective process.

6. **Demonstrating the universal nature of stress** would also serve to defend my definition of stress as an emotion, because basic emotions are common across all humans (Ekman, 1992; Panksepp, 2005). Thus it follows that, as functional state with equivalence to affect, stress should also be present in people from different cultures and countries around the globe. However, in practice this is very difficult to demonstrate as cross-cultural studies in stress research are confounded conceptually in three key areas (Laungani, 1996); 1) many cultures lack an appropriate vocabulary equivalent for stress that captures the same meaning as the Western notion of stress\(^{24}\), 2) the conceptual schema in which stress is defined tend to be very pro-Western, 3) instruments for measuring stress tend to be of Western design and, as “psychometric tests reflect the underlying values of the culture in which they were constructed... they are cultural products” (Laungani, 2002, pg141), their employment in cross-cultural studies tends to invalidate any conclusions such a study would put forward. However, despite these (significant) limitations, work in the study of culturally-specific trauma in India does support the notion that stress is a universal process (Laungani, 2002), even in collectivist cultures. Less rigorous studies (using Western instruments) also support this notion as equivalent stress concepts have been documented in

---

\(^{24}\) This is particularly true for cultures where collectivism is valued over individualism.
Mexican (Mirowski & Ross, 1984), Chinese (Dyal & Chan, 1985), Korean (Klassen, 2010), Greek (Georgas & Giakoumaki, 1988), Russian (Poltavski & Ferraro, 2003) and Jordanian cultures (Hattar-Pollara & Dawani, 2006). Additionally, stress has also been identified across many different cultures through meta-analyses of well-being (e.g. Fischer & Boer [2011] who identify stress in studies across 63 different countries). Thus I posit that stress is universal and, therefore, that it demonstrates concordance with my definition of emotion.

7. **Stress follows the same temporal profile as other emotions.** As stated above, emotions are frequently characterised as having an abrupt onset, which fades away over time (Frijda et al., 1991; Ekman, 1992). This is also true for acute stress, as brief stressor exposure elicits immediate activation of adrenal chromaffin cells via the SAM axis, resulting in dramatic elevation of catecholamine release above tonic levels and also the release of a host of other circulatory neuropeptides as well (Fulop et al., 2005). Whilst catecholamine secretion produces an immediate adaptive effect (e.g. increasing heart rate), activation of the HPA axis and other adrenomedullary neuroendocrine systems allow for physiological responses to be graded, subtle and 'primitively specific' (Goldstein, 2010) - i.e. by varying patterns of neuroendocrine activity (Goldstein, 2003) even the immediate, automatic response to stressors can be stimulus-specific and capable of adapting physiology to support the cognitive responses to the stressor - precisely the same role as that conceived for the emotions.

8. **Studies on animal welfare** often show that animals experience phenomena similar to human stress. This can seen in experiments using
stressor-induced physiological biomarkers of stress (e.g. see Borell et al., 2007), or through endocrine responses to stressor-exposure (Mormède et al., 2007). As animals also suffer from similar stress-related diseases to humans (Blokhuis et al., 1998), some authors have suggested that the term ‘welfare’ ought to be reconceived in terms of AL and stress (Korte et al., 2007). Therefore, at least in response-based terms, it seems likely that animals experience stress on par with humans.

However, research from other studies suggests that animals also experience basic emotions like fear (Forkman et al., 2007; see also Panksepp, 2005), which are conceived to hold similar adaptive value to that of human emotions (Fraser & Duncan, 1998) and, most researchers agree, are initiated as a consequence of basic cognitive appraisal\(^{25}\) (e.g. Ursin & Erikson, 2004). This has led some authors to advocate an affect-based definition of welfare (Dawkins, 1980; Duncan, 1993). According to this view, animal stress overlaps considerably with animal emotion and, I therefore argue, can be considered as one and the same.

9. A final source of evidence for considering stress as an emotion is that “both literatures prominently share the common theoretical framework of appraisal theory” (Smith & Kirby, 2011, pg195). Appraisals form one of the major reoccurring themes across models of affect. Most theories tend to divide appraisals into those that are cognitive and those that are preconscious, or automatic. Although some theories do not view appraisal as the eliciting mechanism of emotion (e.g. Russell’s [2003] Core Affect Theory); nevertheless, the significant majority of theories of emotion do

---

\(^{25}\)Which, according to Erikson, is a phylogenic faculty germane to all vertebrates (Erikson et al., 2005).
include a role for the emotion-antecedent processing of stimuli at some level (Smith & Kirby, 2000; Scherer, 2009; Roseman, 1996; Lazarus, 1999), be it cognitive, automatic or multi-level (Moors, 2009).

**Cognitive Appraisals** are difficult to define and their exact nature is very much open to discussion (e.g. Ellsworth & Smith, 1988). For example, Lazarus views cognitive appraisals as "the process of categorising an encounter, and its various facets, with respect to its significance for well-being" (Lazarus & Folkman, 1984, pg31), which is very similar to Scherer's definition, where appraisals are seen as "the evaluation of the relevance of environmental stimuli or events for the organism's needs" (Scherer, 1982, pg556). Thus both definitions see the key factor in cognitive appraisal as the perceived relationship between the individual and his or her environment. Although this view has been criticised (e.g. Manstead and Fischer [2001], who argue that cognitive appraisal must consider social context; or Parkinson [2001], who believes environmental context is important in the appraisal process) the majority of structural models of cognitive appraisal (e.g. Lazarus, 1991; Roseman 1984; Scherer 2009; Smith & Ellsworth 1985; Smith & Lazarus 1990; see also Scherer, 1988; Schorr, 2001) are in common agreement that cognitive appraisals constitute:

"Some sort of evaluation of how important or relevant the stimulus situation is to the person, whether it is desirable or undesirable, [and] whether, and to what

---

degree, the person is able to cope with the situation" (Smith & Kirby, 2001, pg197).

Most authors (see Barrett et al., 2007) also agree that cognitive appraisals should be considered as an on-going or continuous 'hot' process that focuses on personal well-being (Lazarus & Folkman, 1984) and occurs largely outside of our awareness (Kappas, 2001). In this view, cognitive appraisals can either generate emotion directly, or serve (via top-down input) to regulate pre-existing activity in underlying basic emotional circuits "like a set of switches which, when configured in certain patterns, trigger [subcortical] biologically basic emotional responses" (Gross & Barrett, 2011, pg10; see also Gross & Thompson, 2007).

A wide body of evidence supports the assertion that emotions are related to patterns of cognitive appraisal. For example, Smith and Ellsworth (1985) asked participants to recall past emotional events and, using structured interviews, they identified six common cognitive categories of appraisal across participants (pleasantness, anticipated effort, certainty, attentional activity, self-other responsibility/control, and situational control). This finding has also been replicated by others (Shaver et al., 1987; Roseman, 1991; Smith et al., 1993; Smith & Lazarus, 1993a; Scherer, 1997; Kuppens et al., 2003; Tong et al., 2007; see also Lazarus,

27 The exact sequence of cognitive appraisal varies with different theories. Some advocate a flexible order of appraisals (Lazarus, 1999), or a cyclical cycles process (Smith & Lazarus, 1993a). Others maintain that some appraisals (Ellsworth, 1991) or all (Scherer, 2001) occur in a fixed order.

28 The 'heat' refers to emotion-generation, which functionally separates appraisals from 'cold' factual or situational construals (Weiner, 1985), which provide the underlying knowledge required to support cognitive appraisals, but do not contribute in the generation of the emotion itself (Smith et al., 1993).
Recent evidence from neuroimaging studies also implicates the role of cortical areas in appraisal. For example, Drevets and Raichle (1998) observed increases in blood flow to neocortical areas (specifically the PFC and ACC\textsuperscript{29}) during cognitive emotional tasks. Similarly, Lane and colleagues (Lane et al., 1997a; 1997b) found increases in neocortical activity in PET scans when participants were shown positively- and negatively-valenced photographs. Other studies also advance cogent evidence for a cognitive cortical appraisal system capable of generating and regulating emotion in response to stimuli (Reiman et al., 1997; Mobbs et al., 2007; see also Ochsner & Gross, 1998a).

However, despite this supporting evidence (and general agreement) that cognitive appraisal processes are capable of eliciting emotion, progress establishing the precise nature of these appraisals has proven slow\textsuperscript{30} and the chief criticism of cognitive appraisal theories remains their lack of specificity\textsuperscript{31,32}, either in terms of the precise cognitive neurological circuitry involved in appraisal (although see Ochsner & Barrett, 2001), or in terms of the individual dimensions of appraisal, which vary considerably from researcher to researcher (e.g., Roseman et al., 1990; 1996; Scherer, 2001).

\textsuperscript{29} Although their study also showed decreases in flow to these areas when attentionally-demanding tasks were employed.

\textsuperscript{30} This is commonly ascribed to methodological limitations in assessing cognitive appraisal (Siemer et al., 2007).

\textsuperscript{31} This causes two further difficulties for appraisal theory; 1) it makes appraisal theory very difficult to falsify and, 2) apparently similar appraisal patterns are capable of eliciting very different emotions in different individuals (Frijda & Zeelenberg, 2001).

\textsuperscript{32} Further criticisms of appraisal theories are: 1) that most studies supporting cognitive appraisals use correlation analyses (Parkinson & Manstead, 1992; 1993) which, in addition to their failure to demonstrate causality, often fail to consider alternative explanations through their limited methodology (Reisenzein, 1995; Parkinson, 1997); 2) a large number of studies use recall of events, rather than studying the events themselves; 3) often, appraisals seem to be caused by emotions, rather than the other way around (Frijda, 1993; Zillman, 1983; although see Roseman & Edvokas, 2004; Siemer et al., 2007; Smith & Kirby, 2009).
Briefly, models of cognitive appraisal can be roughly divided into two groups; those who recognise a limited number of appraisal dimensions (e.g. Lazarus, 1994; Ellsworth, 1994; Scherer, 1997; see also Ellsworth & Scherer, 2003) and those who believe the number is, effectively, unlimited (Ortony et al., 1988). As described earlier, my view is that the neocortical areas that administrate the cognitive appraisal processes do so by regulating underlying basic emotion systems (Davidson et al., 2000; Ochsner & Gross, 2005; 2008). Therefore, whilst I recognise that a very wide range of cognitive processes may contribute towards appraisal, I conjecture that their collective effect is channelled into the regulation of a finite number of core subcortical emotional systems, which generate the affective state.

Preconscious appraisals are frequently identified as playing a significant role in underpinning emotion (Ekman, 1999; LeDoux, 1998; Teasdale, 1999; Bargh & Williams, 2007) - "evidence to date strongly supports the notion that emotional significance is rapidly and preconsciously detected" (Compton, 2003, pg82). Contrasting conscious and involuntary appraisal systems, Mauss states that, whereas cognitive appraisal processes;

"Require attentional resources, are volitional, and are driven by explicit goals..., automatic processing is initiated by the simple registration of sensory inputs, which in turn activate knowledge structures (schemas, scripts, or concepts) that then shape other psychological functions" (Mauss et al., 2007, pg148).
Beginning with Zajonc's (1980; Murphy & Zajonc, 1983) early discovery that stimuli could elicit emotions automatically\textsuperscript{33}, research has consistently reported that stimuli can not only be attended automatically, but that a significant degree of *processing* and *evaluation* may also take place preconsciously (Bargh, 1990; Bargh et al. 1993; 1996\textsuperscript{34}). If accurate, such processing is likely to be of significant adaptive advantage as, in theory, it facilitates rapid, parallel processing (Robinson, 1998) of a great deal of information (Kitayama, 1990) with little to no attentional resources (Schneider et al., 1984) and, because of its assumed adaptive value, preconscious appraisal may well prove to be significantly larger in capacity than conscious processing (Öhman, 1994; Öhman et al., 1993). Accordingly, a number of theories of emotion include a role for preconscious appraisal (e.g. Clore & Ortony, 2000; Power & Dalgleish, 1999; Smith & Kirby, 2001; Teasdale, 1999).

Initially, the role of preconscious appraisal was viewed either as a stimulus valence-detection process (Robinson, 1998), or in terms of directing the focus of cognitive attention. This is because early research in this field found that patients with emotional disturbances tended to

\textsuperscript{33} A process he ascribed to classical conditioning through repeated stimulus exposure (Zajonc, 2000; 2001).

\textsuperscript{34} In general, experiments demonstrating preconscious evaluation of stimuli tend to make use of either the phenomena of affective priming (e.g. Fazio et al., 1986; De Houwer & Eelen, 1998; Winkielman et al., 1997) or of implicit association (Greenwald et al., 1998; Greenwald & Farnham, 2000). This latter technique has been used frequently to demonstrate preconscious racial prejudice (e.g. McConnell & Leibold, 2001) and other preconscious emotional appraisal processes (Mauss et al., 2006). These studies have been validated through the use of neuroimaging techniques, which have shown conclusively that these implicit association methodologies activate subcortical affective systems to conduct preconscious appraisals. For example, Phelps and colleagues (Phelps et al., 2000) used the implicit association paradigm to show that implicit tasks involving racial stereotyping are associated with amygdala activity. Similar observations have been made by others (e.g. Cunningham et al., 2003), who demonstrate that, whilst explicit (cognitive) appraisals are associated with PFC activity, implicit appraisals are not.
demonstrate an unconscious preference for information relevant to their conditions (Williams et al., 1996), and people with high trait anxiety often show an automatic bias towards information associated with their anxieties (Wells & Matthews, 1994). This led Kitayama (1990) to suggest that the primary function of preconscious appraisals is to act like an unconscious 'spotlight' to direct conscious attention towards high-valence stimuli. Work on the neural basis of fear has now conclusively demonstrated that one of the functions of preconscious appraisals is exactly that - to shift cognitive attention onto dangerous stimuli (LeDoux, 1998; LeDoux & Phelps, 2008) and evidence is growing to support the existence of similar preconscious mechanisms for the other basic emotions as well (e.g. Both’s work on automatic appraisal of sexual stimuli [Both et al., 2011]; see also research into the evaluation of social stimuli [Adolphs, 1998], contextual stimuli [LeDoux & Phelps, 2008], facial expression [Adolphs et al., 1999] and the modulation of long-term memory storage [McGaugh, 2004]). Thus, I advance that one of the primary functions of preconscious appraisals is to evaluate and encode stimulus valence (Bargh & Williams, 2007) and to direct cognitive attention selectively towards high-valence stimuli.

In close alignment with theories of emotion, stress research also recognises the crucial role of stimulus appraisal; indeed, considerable experimental evidence supports this conclusion: Lazarus (1999), for example, details over forty separate studies (many of which have been replicated recently e.g. Compass et al., 2006; Yi et al., 2005) demonstrating the importance of appraisals in eliciting stress. This evincing research is further supported by evidence from physiological studies, which have repeatedly concluded that stressor appraisal
processes are associated with increases in glucocorticoid secretion (Ursin & Eriksen, 2004; Ursin, 1998; see also Dickerson & Kemeny, 2004) and cardiovascular arousal (Brownley et al., 2000) which, as a consequence of threat-type stressor appraisals, have been strongly implicated in the development of stress-related illnesses, including cardiovascular disease (Ohlin et al., 2004), depression (Kessler, 1997), anxiety disorders (Stanford et al., 1993) and accelerated ageing (O'Donovan et al., 2012). Thus, in tandem with emotion theory, I assert that appraisal processes also play a critical role in the generation of stress.

Traditionally, theories pertaining to the evaluation of stressors have coalesced onto two different types of appraisal process; Primary and Secondary (Lazarus & Launier, 1978). Primary appraisals are defined as assessing whether “what is happening is relevant to one’s values, goal commitments, beliefs about self and world, and situational intents” (Lazarus, 1999, pg92) - i.e. an evaluation of “what is at stake in the encounter” (Lazarus & Folkman, 1984, pg33) or, simply put, whether a stimulus is irrelevant, benign/positive or a potential stressor. By contrast, secondary appraisals determined whether a stressor was evaluated as harmful, threatening or a challenge (Lazarus & Launier, 1978) depending on whether the individual believed they had the resources to cope with the stimulus (in which case it was evaluated as 'challenging') or not (when it was appraised as a 'threat'). However, these definitions have increasingly fallen out of use in recent years, chiefly because they appeared to define appraisals as exclusively 'cognitive' events (Zajonc 2001; 1984). Without wishing to re-open the infamous Zajonc-Lazarus debate about the fundamental nature of the stressor appraisal process; most stress theorists, including Lazarus himself, now recognise that
appraisals can be automatic (Frijda, 1986; Smith & Kirby, 2001; see also Bargh et al., 2010) and, mirroring the appraisal of emotive stimuli, frequently occur preconsciously — i.e. entirely outside of conscious awareness (Kappas, 2006). This acknowledgement of the role preconscious appraisal processes can play in the development of the stress process ultimately engendered a reconceptualisation of stress from a multi-level, or dual process perspective, which I will briefly outline now.

**Dual process models of stress**

Emotions, as I have defined them, already share many significant elements with models of stress, leading some researchers to redefine psychophysiological stress responses in affective terms as ‘emotion regulation under stress’ (Compas, 2009; see also Eisenberg et al., 1997; Oschner, 2007; Skinner & Zimmer-Gembeck, 2007). For example, as with affective research, theories of stress have converged on the notion of separate levels of stressor appraisal. Such proposals are collectively known as dual process models of stress (Compas, 2004; Compas et al., 2001; see also Eisenberg et al., 1997; Skinner & Zimmer-Gembeck, 2007). Dual process models identify two distinct types of appraisal process: 1) those that are automatic — i.e. they:

“Occur without intention [often outside of awareness], require relatively little in terms of cognitive resources, and are often difficult to interrupt or terminate once they have been initiated” (Sherman et al., 2008, pg315)

And, 2) a set of controlled, volitional responses to stress, which;
"Require cognitive resources, can be volitionally interrupted or stopped, and
typically [occur] within awareness" (Sherman et al., 2008, pg315).

Evidence for dual process models of stress is sizeable (for a comprehensive summary of this evidence, see Compas et al., 2009) and can be collated into 8 distinct categories; a) research on associative conditioning and learning (Compas et al., 2001), b) studies on strategic-controlled and automatic cognitive processes in emotional disorders (Compas et al., 2001; Compas et al., 2009), c) research into stress-induced changes in attention (Bar-Haim et al., 2007), d) research differentiating temperamental characteristics from intentional, cognitive processes (Compas et al., 2001), e) research on automaticity in social cognition (Matthews & Macleod, 1994), f) neuroimaging studies (Gianaros & O'Connor, 2011), g) studies analysing stress-induced manipulation of memory (Carver & Cluver, 2009), and h) evidence from research into stress-induced alteration in executive functioning (Compas, 2006). In addition, dual process theories have also been independently corroborated using self-report questionnaires from adolescents and their parents (Connor-Smith et al., 2000) and through separate, confirmatory factor analyses studying a large range of stressors in a diverse variety of people (e.g. Connor-Smith & Calvete, 2004; Wadsworth et al., 2004).

By dividing stressor appraisals into two levels, dual process models of stress separate psychophysiological responses to stressors functionally. Preconscious responses tend to describe automatic appraisal processes in terms of stress reactivity (automatic, rapid responses to stressors), which includes stressor valence detection processes (Bargh & Williams, 2007), allocation of attentional resources (Williams et al., 1998; see also
Bar-Haim et al., 2007) and, possibly, implicit activation of long-term memory and goal-dependent behaviour (Bargh et al., 2012). By defining these "immediate and automatic reactions to stressful situations" (Compas et al., 1999, pg231) as 'action tendencies' (Skinner, 1999) - i.e. processes which amalgamate motor programmes and core affect (e.g. fear or shock) with goal orientation (Skinner & Zimmer-Gembeck, 2007) - Skinner is, effectively, equating the role of these involuntary stress responses with the core processor components of basic emotions (Panksepp 1998; Ekman & Cordaro, 2011), to which I have already ascribed responsibility for initiating very similar 'states of "action readiness" (Frijda & Parrott, 2011). Indeed, the similarity is further emphasised by examining the adaptive roles of these processes, which are both clearly defined in terms of efficient, rapid adaptation to environmental stimuli. Additionally, as neuroimaging studies (e.g. McEwen & Gianaros, 2010; Pruessner et al., 2010) tend to place preconscious stressor appraisal processes in exactly the same subcortical areas (particularly the hippocampus and the amygdala - Davis & Whalen, 2001) as multi-level models of affect, this serves clearly to emphasise my assertion that stress and emotion are functionally synonymous; particularly in terms of the neurobiological operation of their respective preconscious responses.

In addition to the automatic responses to stressors, dual process models of stress also identify an over-laying set of volitional, conscious processes, which "are captured by the concept of coping," (Compas, 2009, pg93; see also Skinner & Zimmer-Gembeck, 2009). These systems are purported to mediate the longer-term response to stressors and are also reported to be capable of significantly more diverse regulation of the stress process than the automatic 'action tendencies' (Skinner, 1999). As
with the preconscious processes, neuroimaging studies frequently locate coping-based systems within neocortical brain regions (particularly the ACC, PFC and Cingulate Cortex - Gianaros & O’Connor, 2011) heavily implicated in the regulation of emotion (Damasio, 1994), leading Bruce Compas to suggest that coping and emotion regulation could be unified within one construct (Compas et al., 2009)\textsuperscript{35}. Additionally, and in direct parallel with emotion regulation, because coping responses are “shaped by the resources and contexts in which they unfold, [they] are virtually infinite in their variety” (Skinner & Zimmer-Gembeck, 2007, pg124) and thus it has been proposed that a very wide range of cognitive processes may contribute towards the ‘regulation of stress’, or coping (Compas, 2009) - a system that draws many parallels with the top-down contribution to emotion I reviewed earlier in this Chapter.

Although there is some disagreement (e.g. Cole et al., 2004), most researchers accede that the conscious and preconscious components of dual process theories are likely to function either simultaneously (Campos et al., 2004), or in parallel (Compas et al., 1999) with a significant degree of mutual influence (Skinner, 1999; Compas et al., 2001; 2007). Whilst I suggest that this consensus appears very similar to the quorum of opinion within emotion regulation research, nevertheless, some interesting differences have been documented between the two constructs. Firstly, coping is specific to the regulation of stressors, whereas emotion regulation incorporates responses to a much wider range of stimuli (Aldwin et al., 2011). Secondly, and in tandem with the first point, “coping is an organisational construct, encompassing the regulation of multiple

\textsuperscript{35} Compas’ suggestion was that executive function could unify coping and emotion regulation.
processes” (Skinner & Zimmer-Gembeck, 2009, pg7); whereas emotion regulation, by contrast, appears limited to the control of behaviours, responses and systems associated with affect. Should dual process models of stress and emotion regulation become unified as Compas (Compas et al., 2009) and I both advocate, I suggest that the differences outlined here between the concepts would support a role for emotion regulation to become subsumed within the more expansive concept of coping. Although further work is required to explore this relationship, it is promising to note that recent new research (e.g. Wang & Saudino, 2011) has already begun to provide evidence which not only supports such a relationship, but also underscores the commonalities I have previously described that appear to exist within these literatures.

The final difference that appears to exist between emotion regulation and dual process theories of stress is that, within stress research, the majority of investigations have focused on the role of descending neocortical input in regulating preconscious systems (e.g. see Compas, 2006) - the ‘top-down’ approach - chiefly because it has been reported to exert a significant effect on the allocation of conscious attention to stressors (Mogg et al., 2004; Pine et al., 2005) as well as other forms of secondary control coping, such as disengagement from stressors (Copeland & Compas, 2006) and stressor reappraisal (Compas, 2009). By contrast, as I explained earlier in the Chapter, within emotion regulation research, some considerable evidence upholds a role for regulation in both directions (i.e. top-down and bottom-up). This difference is surprising on two grounds; 1) because it is very much out-of-keeping with my assertion that emotion regulation and the stress process should be considered as functionally equivalent, and 2) because, as I explain more
comprehensively in Chapter 6, it is incongruous with findings from studies of illnesses like anxiety, addiction and depression, where the influence of preconscious processes is believed to have particular significance in the development of psychopathology. Taken in tandem, these two observations highlight a potential 'gap' in stress research - i.e. the regulation of acute and chronic stress via changes to preconscious mechanisms. It is this area of research that I wish to explore further in this study, by examining the extent to which preconscious appraisals change during the acute stress process and, as I explain in the next Chapter, whether specific cognitive resources well-established in volitional stress regulation, could also play a role in preconscious stressor appraisal.

**Stress as an emotion - a synopsis**

Thus far I have presented an overview of emotion from a neurobiological, multi-level perspective. I have also outlined nine sources of evidence supporting my view that dual process models of stress share functional equivalence with multi-level models of emotion. Before advancing any further arguments I wish briefly to summarise the ramifications of my view.

To conclude, by viewing stress as an emotion I assert that;

1. Stress is an evolutionarily adaptive process that prepares individuals physiologically and neurologically to deal with exposure to stressors.
2. Stress may also become maladaptive if stressor exposure is chronic and/or if stress leads to physiological and/or neurological pathologies.

3. Stress is generated by the activity of subcortical emotional systems, either directly or via top-down neocortical activity. The action of stress hormones upon these systems also affects stress further.

4. Stressors are identified as such through two separate, but interrelated processes – namely cognitive appraisals and preconscious appraisals.

5. Both appraisal processes are capable of eliciting stress.

6. The evaluation and subsequent identification of a stressor as such by cognitive appraisal is affected by a wide range of goals, beliefs, knowledge, desires, emotions, memories and cognitions.

7. Preconscious processes may also draw on many (or possibly all) of these mental representations to appraise stressor stimuli.

8. Both appraisal processes are capable of self-regulation via a range of antecedent- and response-focused strategies.

9. Both appraisal processes are capable of regulating or changing each other, either directly or indirectly.

Having presented by argument that stress possesses functional equivalence with affect, in the next Chapter I move on to examine the impact of this position on social cognitive theories of human agency. I then explore the extent to which efficacy beliefs may interact with the stress process and propose three Research Questions to develop the ‘gap’ I identified earlier and thereby investigate the role preconscious stressor appraisals play in the stress process.
Chapter 3: Humans as Efficacious Agents

Introduction

Social cognitive theories of personality and motivation hold at their centre the view that people are emergent and interactive agents (Bandura, 1977; 1989; 1997): we are neither autonomous, nor simple mechanical conveyors of animating environmental influences. "Rather, we make a causal contribution to our own motivation and action" (Bandura, 1989, pg1175) via our "beliefs, values and goals, which set up a meaning system within which we define ourselves (Dweck, 2000, pg138-139). Thus, to be an agent is to influence intentionally one's functioning and life circumstances:

"People are self-organizing, proactive, self-regulating, and self-reflecting. They are not simply onlookers of their behaviour. They are contributors to their life circumstances, not just products of them" (Bandura, 2006a, pg164).

Bandura, therefore, presents agency in functional terms - "to be an agent is to influence the course of events by one’s action" (Bandura, 2011, pg52). A variety of other conceptions of agency exist (e.g. Blakemore & Frith, 2003), of which, I feel, Pajares' definition comes closest to summarising a specific definition of agency as Bandura views it. Pajares states that agency is:

"The sense that, among other personal factors, individuals possess self-beliefs that enable them to exercise a measure of control over their thoughts, feelings, and actions" (Pajares, 2002).
Central to this definition of agency, (and implicit in Bandura’s [1997; 2006a; 2008] ‘core properties of agency’ – see later) is the concept of ‘self’. ‘Self’ is a diffuse term and, accordingly, its definition presents something of a challenge! However, despite cross-cultural variations in the perception of ‘self’, Bukobza presents the consensus view that ‘self’ is ‘a distinct and independent unit’ germane to all humans (Bukobza, 2007) and that, in contrast to transcendental views, ‘self’ is generated in the brain (Churchland, 2002). This view is corroborated by evidence from split-brain studies (Sperry, 1982) and a large body of research into self-related clinical disorders (Feinberg & Keenan, 2005). However, despite numerous neuroimaging studies (e.g. Gusnard, 2005; Gillham & Farah, 2005; Northoff et al., 2006), convergence on a single neurological location for ‘self’ has proven difficult. There are a number of reasons for this impasse which, I suggest, could owe to limitations in neuroimaging methodologies (Christoff et al., 2011), the transient nature of ‘self’ (Strawson, 1999), multiple neurological locations for ‘self’ (Feinberg, 2001; Feinberg & Keenan, 2005) or that ‘self’ may be an emergent phenomenon and, therefore, “different from and not reducible to neural events” (Sperry, 1977, cited in Trevarthen, 1990, pg384). As Sperry’s position appears similar to Bandura’s (1997; 2006a; 2006b; 2008), I address the issue of reducibility later in this Chapter. However, before doing that, I wish first to set down my own position on the nature of ‘self’, as I believe a clear characterisation of ‘self’ is critical in establishing an unambiguous concept of human agency.
The neurological ‘self’

LeDoux’s (2003) suggestion of a multifaceted ‘self’ located across a number of brain systems seems to me to fit more pieces of the puzzle together than other suggestions about the nature of ‘self’. Firstly, this position adopts the ontologically reductive position of ‘grounding the self’ (Damasio, 2000) in the brain, which has become a widely accepted view\(^{36}\), even within the traditionally functionalist disciplines of psychology and cognitive science - “questions about self-representation are steadily shifting into the province of the brain and cognitive sciences” (Churchland, 2002, pg309). Secondly, LeDoux’s model of ‘self’ requires no single neurological ‘locus’ for ‘self’\(^{37}\), which explains the problems neuroimaging studies have encountered finding one. Thirdly, it begins to account for the multitude of different tasks and responsibilities people ascribe to their concept of ‘self’, which would certainly require, at the very least, distributed access to (and, according to LeDoux, integration with) many different brain areas. Fourthly, it encompasses multi-level interaction between ‘self’ and affect, possibly via a ‘shared neurobiological stem’ (Damasio, 2003a; see also LeDoux, 1998), which accounts for the repeatedly observed link between self-processing and emotion (Fossati et al., 2003). Fifthly, it accepts the notion that, at many levels, ‘self’ is subconscious (Gallagher, 2000), which is a difficult conclusion to avoid when considering self-recognition studies in neonates (Meltzoff & Moore, 1977; 1983) and higher mammals (Beshkar, 2008). Finally, it also implicates ‘self’ in the generation of consciousness

---

\(^{36}\) This is also Bandura’s stated opinion (see next section).

\(^{37}\) Or any equivalent Cartesian homunculus.
Within this multifaceted 'self', I find Damasio's (2000; 2003b) proposal of a protoself most compelling. Similar in concept to Panksepp's Simple Ego-type Life Form (SELF - Panksepp, 1998; 2005), both 'self'-models constitute an internal body map, which functions both as a core representation of 'self' and also as a reference point for the generation of other self-related systems (Damasio, 2003b). Damasio (2000; 2003a) advocates the existence of the protoself on the grounds that, a) the body is relatively static and, b) evidence exists to suggest that the brain continually represents the structure of the body internally (see Craig, 2002). This 'internal representation' was first documented in the 1950s, when Holst, Mittelstaedt and Sperry all observed independently that central motor commands are also accompanied by an identical, redundant, parallel copy command. Whilst Damasio (2000) and Panksepp (1988) use this as the starting point for their nuclear 'self' models, Frith (1992) suggested that this 'efference feedback system' could also be used as the neural basis of human agency.

According to Frith, the 'efference feedback system' generates self-awareness of our own actions. He cited a number of studies into somatoparaphrenia - 'alien hand syndrome' - as examples of the system

---

38 Although note here that LeDoux (1998) and Damasio (2003b) have very different conceptions for the generation of consciousness. LeDoux believes it was layered on top of the brain by recent evolutionary processes, whereas Damasio's proposal establishes rudimentary consciousness from the beginning of evolution.

39 Although note that Damasio and Panksepp suggested different brain areas for the basis of their models.

functioning improperly - i.e. when our own movements are not associated as 'self'. The opposite observation can be seen in anosognosic patients, who believe they have initiated movement when, in fact, they have not. In this case the motor command is believed to fail, whilst the efference feedback copy records the movement as having taken place (Heilman et al., 1988). Frith (1992) suggested that a centrally-driven comparison of the efference command copy and the motor command could account for the sense of agency as it would produce a self-responsibility 'label' for our actions. Frith also used this as a model for schizophrenia: schizophrenics, he argued, were self-aware of their movements but, through a failure of the comparator process, did not experience a normal sense of agency. Whilst Frith's argument is, at many levels, compelling, Vignemont and Fourneret (2004) point out a number of scenarios where this explanation cannot fully account for agency (e.g. in paranormal phenomena such as automatic writing, where a conscious system must also contribute to the sense [or absence] of agency). Thus, whilst the 'self' map advocated by Damasio and Panksepp is a very helpful starting point, I conjecture that it cannot account for the full sense of agency, or of 'self'. In order to do that, I need to recruit other aspects of LeDoux's multifaceted self.

Markus and Kitayama (1991, pg229) assert that the 'self' must contain "self-relevant schemata, [which are] used to evaluate, organize, and regulate one's experience and action". As I explain later, the key idea here is experience, or memory. At first sight, LeDoux's multifaceted self may appear analogous to Dennet's (1989; 1992) *Joycean Machine*, in which a decentralised model of cognition drives a human machine, much like a 'self-steering ship' (Dennet, 1992). In this system, 'self' is illusory - a non-
causal by-product of the steering mechanism. However, what separates Dennet's 'Joycean Machine' from the multifaceted 'self' is Markus and Kitayama's (1991) observation that experience is crucial in generating the subjective, or autobiographical (Damasio, 2000) nature of 'self'. As Ismael's (2006) rightly points out, the Joycean Machine model would predict history encoded as an objective record of events, rather than the autobiographical, subjective (and, frankly, patchy) narrative that is actually remembered. Thus, the multifaceted 'self' does not function as a Joycean Machine because it comprises an autobiographical 'self' in addition to the 'steering mechanisms' of the protoself. Gallagher (2000) also advocates the existence of 'self' in a narrative form - a “self extended in time to include memories of the past and intentions toward the future” (Gallagher, 2000, pg15). Both Damasio and Gallagher base their experiential forms of 'self' within episodic memory systems, which theoretically enable one to construct a 'narrative' of 'self' over time and also to project that 'self' forward in time. However, as "episodic memory is constantly remodelled by innate and acquired dispositions as well as social and cultural environments" (Gallagher, 2000, pg20), it follows that the ability to self-recognise in episodic memory requires a minimal and constantly reiterated sense of self, which Damasio (2000) argues comes from reference to the protoself generated via the efference feedback system. Whilst evidence from amnesiacs seems to refute the idea of a narrative self (Churchland, 2002), studies with patients suffering from other clinical memory disorders reliably supports the suggestion of a memory-based 'self' as disordered memory commonly affects self-

41 Dennet draws the parallel of the mind as a termite mount, where centralised control appears to be present, but is actually an illusion generated as a by-product of a self-organising system of autonomous parts.

42 A pivotal role in the development of self-belief systems such as SE.
concept (Schacter et al., 2003). Therefore, within the neurobiological mechanistic model of 'self' I have sketched out here, I conceive that there are two key conceptual components:

1. A narrative, subjective 'self', which functions as a context-dependent, subjective process (Damasio, 2003a; 2003b) that is bi-directionally integrated with conscious and unconscious affect regulation processes, cognitive cortical areas and memory systems.

2. An underlying protoself (Damasio, 2000) or SELF concept (Panksepp, 1998), which acts as a source of reference for the higher system/s.

Thus, similar to LeDoux's (2003) multifaceted view of 'self', I posit that 'self' is best understood from a mechanistic neurobiological position (Craver, 2006 - see next section) such that 'self' is produced by a 'non-nested hierarchy' of brain systems (Feinberg, 2001; Feinberg & Keenan, 2005) which generate 'self' as a collation of mechanisms. This 'self' is not a Cartesian 'I' (as it is influenced by unconscious and affective 'bottom-up' processing), but neither is it sensory-driven 'mindless automaton' (Bandura, 2006a). It is a unified orchestra of systems, whose processes assemble to generate a unique and biologically-contextualised 'self', whose experience and subjective monologue establishes a critical foundation underlying human agency.

---

43 Their work on implicit memory also supports the role of 'self' generation via subconscious processing.
On reductivism

For Bandura, human agency is founded on the principle of 'non-reductive physicalism' (Bandura, 2008; 2006b; 1989) and, as such, he defines agency as an ontologically emergent, materialist phenomenon (Bandura, 1991b; 2006b) which, despite its dependence on underlying brain processes, is not epistemologically or methodologically reducible to them - i.e. agency is derived from neurological processes in the brain, but cannot be explained by studying them at their component levels. I contest this position and, whilst I firmly subscribe to Bandura's view that agency plays a central causative role in governing human behaviour, I do not adopt his definition of agency. I view human agency from an ontologically reductive neurobiological perspective which, because it is at odds to Bandura's view, necessitates a brief critique of Bandura's anti-reductive stance as applied to human agency.

In many of his recent papers, Bandura (2008; 2006a; 2006b) is critical of biological models of agency, which he suggests have a tendency to be dehumanising due to their reductionist underpinning:

"Self-regulatory processes are being dismissed in some quarters on the grounds that human behaviour is regulated unconsciously by neural networks with illusory belief that one is exercising personal control. This type of conception strips humans of agentic capabilities, a functional consciousness, and a self-identity" (Bandura, 2006b, pg3).
This argument has been repeated in subtly different guises by many other authors. For example, Melden (1961) presents the challenge of the 'disappearing agent' such that biological-behaviouristic processes reduce the agent to "a helpless victim of the conditions in my body and its immediate physical environment" (Melden, 1961, pg 129). Whilst Schlosser (2008) presents an excellent counter-argument to the reductionist dilemma of the 'disappearing agent' (which I outline in a footnote), Bandura's position appears to maintain that reductionist philosophy falls short of encapsulating a meaningful concept of agency as he views it.

Theoretical reductionism, such as that put forward by Nagel and Schnaffer (Nagel, 1986), traditionally focuses specifically on ontological, or theoretical reduction i.e. processes that derive reduced laws from reducing theory. This approach has been heavily criticised (e.g. Sakar, 1992; see also Dijadzi-Bahmani et al., 2010), particularly in biological fields, where philosophers have argued that underlying biological laws simply do not exist (although see Sober [1993], who argues that predictive patterns can be considered as laws), particularly in genetic- and evolutionary processes (Waters, 2004). This is compounded by the

44 See also Nagel (1986), who argues that neurological 'events' must be attributable to other events in a causal chain. According to Nagel, this "leaves no room for action" (and hence agency) as every event is pre-determined by another.

45 Schlosser (2008) argues that hierarchical standard-causal theories of agency allow for the type of rigid behaviourism, or biological reflexivity Melden describes, but also for deliberate-, reflective- and autonomous action, provided the agent is ceded the ability to deliberate between different behavioural outcomes (an effective control system) and a schema of attributes and beliefs with which to evaluate those outcomes. This is not without some irony when one considers that it is precisely this system of beliefs that forms the lynchpin of Bandura's notion of human agency.

46 On par with those of other Natural Sciences, e.g. the laws of thermodynamics. This is particularly the case for genetics and proteomics, where the example of protein folding or genetic coding is often cited (e.g. see Huttenman & Love, 2011 for a full explanation).
dependence of evolutionary mechanisms upon explicit environmental contexts (Hull, 1972: 1979), which also introduces the further problem of multiple realisable outcomes (e.g. convergent evolution) when theories overlap each other (see Sklar, 1993, who cites the example of pain as a non-reducible concept on the grounds of multiple realisability). However, whilst theoretical reductionism may well have significant challenges to face within the critical realist paradigm (and particularly in biological research) it is by no means the only reductive position available (e.g. Dowell, 2006). This is my main criticism of Bandura’s position – he collectively dismisses reductionism en masse without appearing to make any visible account for different reductive epistemologies. For example, explanatory reductive theory\(^{47}\), such as that proposed by Kauffman (1971) in his ‘articulation of the parts’ process, does not adhere to the truth-derivation system central to theoretical reductionist epistemology. Instead, it attempts to establish a causal explanation for the behaviour of a system in terms of the properties of the component parts and the relations between them. Effectively, Kauffman is advocating a form of pragmatic reductionism that allows researchers to make knowledge claims about the respective parts of a system towards the goal of understanding how the system works collectively, but without the requirement that theory dissolve entirely into reduced system components. I posit that such explanatory epistemology may overcome a large number of the problems associated with theoretical reductive approaches, chiefly because it does not hold towards the same stringent ontological requirements about the nature of reducible knowledge.

\(^{47}\) See also Part-Whole approaches (e.g. Winther, 2011)
Mechanistic theory (Craver, 2006) is a subset of explanatory reductive philosophy that has been extensively applied to the study of cognition and neurology (Craver, 2007). Briefly, this approach postulates that the overall behaviour of a system (e.g. the brain) can be decomposed into the activity of mechanisms that underpin the activity. Mechanisms are "entities and activities organised such that they are productive of regular changes from start or setup to finish or termination conditions" (Machamer et al., 2000, pg3) and are themselves defined as constructs of smaller entities (physical structures e.g. cortical lobes or thalamic regions) and activities ("things that entities do" [Craver, 2006, pg371] i.e. causal actions e.g. synaptic action). Whilst some philosophers view mechanistic approaches as purely reductive (e.g. Bickle [2003], who argues that psychological approaches should "lose their initial status as causally-mechanistic explanations" (pg110) when “real neurobiological explanations” are on offer [pg111]), by contrast Craver (2007) argues that mechanisms should be considered non-reductive insomuch as they have no discernible, intractable hierarchy of explanation: “there are no monolithic levels in biology” (Craver & Bechtel, 2007, pg190) - a point they assert is particularly relevant to the study of neurobiology. Craver (2007) cites Hebbian LTP as a good example of non-reductive mechanistic theory in evidence, as the discovery of synaptic potentiation in 1973 has contributed not only towards a 'top-down' reductive search for a molecular model of synaptic memory, but also to extensive 'bottom-up' research into systems of memory formation - another example of multilevel mechanisms in action. As stated previously, Bandura seems to make little consideration of mechanistic models of agency, and I suggest that, for him, a sparse dichotomy exists, which appears to catalogue agential theories into either reductive or non-reductive brackets. I
reject this position on the grounds that multiple, non-reductive perspectives are available (such as Craver's mechanistic perspective) and propose that the concepts of agency and 'self' fit well into the ontologically reductive, neurobiological mechanistic framework I outlined previously. In addition to the criticism previously presented, there are three lesser (but important) considerations which also introduce clear water between my position and Bandura's non-reductive perspective. I outline these arguments as follows.

For Bandura, agency is very much more than the sum of its neurological parts. To make this point, he cites the example of education, where a "thorough knowledge of the brain circuits involved in learning does not tell one much about how best to devise conditions of learning" (Bandura, 1991b, pg157). He advocates the notion of 'second-order control' (Bandura, 2008) to explain his position, arguing that:

"In acting as agents, individuals obviously are neither aware of, nor directly control, their neuronal mechanisms... They intentionally engage in activities at the macrobehavioral level" (Bandura, 1991b, pg160).

Thus, to paraphrase the emergent physicalist argument (Bandura, 2006b) - the brain generates agency, but neurobiological methodologies are insufficient to provide an understanding of agency. Apart from the reservation previously identified, my main criticism of this view is eloquently voiced by Rottschaefer (1991), contends that, by adopting a position of physicalism, despite his assertion that agency is emergent, Bandura is, nevertheless, arguing from an ontological reductionist perspective. Thus his philosophical standpoint is, paradigmatically,
subject to the same epistemic principles as the neurobiological mechanistic view I propose. This is because both standpoints derive from "regularities that are predictive and explanatory of observations [which has] led to the formation of [both] theories explanatory of these regularities" Rottschaefer (1991, pg154). In other words, all physicalist positions (whether Bandura’s or my own) attempt to predict behaviour empirically. Given this, Rottschaefer argues that an application of Occam’s razor would lead to a rejection of Bandura’s notion of agency in favour of a neurobiological approach on the grounds of complexity - it necessitates an additional irreducible construct that neurobiological models do not require. Therefore, in addition to my first argument for rejecting Bandura’s notion of irreducible agency, I also reject it on the grounds of parsimony when a mechanistic position is available.

A second rationale for adopting a mechanistic position over that described by Bandura (2008; 2006b; 1997) is epistemological, rather than ontological. In a general criticism of cognitive theory Eysenck (1989) argues that 'cognitions' themselves are simply metaphors for underlying complex physiological brain processes and, therefore, must be considered only to be "loose descriptions of processes which we do not understand" (Lee, 1992, pg259). Kimble (1989) also highlights the potential danger of confusing such 'descriptive metaphors' with actual explanations of behaviour, warning against the tendency of psychologists to pass off post-hoc concepts as explanations. Lee (1989) draws these reservations together into one specific criticism - Bandura appears not to make an attempt to provide an underlying neurobiological explanation for agency

Note that the principles themselves do not have to be identical merely that they subscribe to the same ontology.
(quite the opposite in fact!) which, I conjecture, effectively makes the concept 'unobservable, unverifiable and unfalsifiable' using post-positivist methodology. A similar argument is advanced by Craver (2006) when comparing representational and phenomenological models. He argues that, although phenomenally accurate models may predict behaviour exceptionally well, they are at heart non-explanatory and, therefore, limited in power. He couches this in terms of 'how-possibly' and 'how-actually' models and argues that "constitutive explanations require descriptions of real mechanisms, not mere how-possibly posits" (Craver, 2006, pg374). Thus, I also reject Bandura's 'how-possibly' concept of irreducible agency in preference to a 'how-actually' neurobiological mechanistic model (even if that model has less predictive power) on the grounds of increased explanatory potential.

In close relation to the previous point, a third criticism of Bandura's position is highlighted by Lee (1992). She points out that, without a testable model of agency, methodologies for data collection are severely limited;

"A great deal of research with a cognitive orientation seems to be no more than the correlation of self-reports of hypothetical constructs with self-reports of other hypothetical constructs" (Lee, 1992, pg259).

As a 'how-actually' model, the neurobiological encapsulation of agency I propose below is, at least theoretically, more testable methodologically than Bandura's irreducible concept. Thus, I assert that my position is, therefore, preferential on the grounds of methodological scope. For example, whilst Bandura's concept of agency is frequently limited
methodologically to self-reports (as Lee [1992] attests), hypothetically a neurobiological model is congruent both with this method of data collection and also with a wide range of other post-positivist empirical methodologies (e.g. neuroimaging, genetic analysis or animal modelling) many of which, by nature of their post-positivist epistemology, make available the principle of proof by experimental falsification (Popper, 1959). I do not wish to critique this philosophical position here (and I readily acknowledge the ontological equivalence of the ‘plurality of post-Khun research perspectives’ [Popkewitz, 1984] now available); rather, my point is that the neurobiological mechanistic model I propose seems to make available a considerably greater range of methodologies than Bandura’s non-reducible concept, at least from a theoretical perspective. Furthermore, as many of these additional methodologies offer alternatives to simple correlational analysis (a major limitation of all self-reports – see Chapter 6), I reassert that the neurobiological mechanistic approach I adopt here may possibly constitute a more pragmatically versatile model for understanding human agency than Bandura’s original concept. I outline my understanding of agency in the next section.

**Humans as efficacious agents**

In the previous sections I asserted that agency could be ontologically reducible to neurobiological brain mechanisms and attempted to explain how these systems, when collated, might generate the 'self' that supports human agency. In this section I explain how neurobiology may account for human agency and I focus on efficacy beliefs as a key component of agentic functioning.
Human agents “exercise a measure of control over their thoughts, feelings and emotions” (Pajares, 2002, pg1). My emphasis in this definition is crucial, because ‘control’ forms a vital lynchpin of Bandura’s conception of agency and also a focus for his critique of neurological models of agency:

“The exercise of human agency is dismissed by physical eliminationists on the grounds that human behaviour is regulated by neuronal mechanisms operating at a subpersonal level outside of one’s awareness and control. Deliberative, reflective, self-referential, and other high-level cognitive events are dismissed as epiphenomenal events that create an illusion of control but actually have no effect on how one behaves. In this view, humans are essentially conscious hosts of automata that dictate their behaviour subpersonally” (Bandura, 2011, pg52).

Although I agree with the principle of human agency, I also hold three reservations the view above: 1) to me, it seems to assume that neurobiologists adopt a de facto view of agency as epiphenomenal, 2) to me, it appears to describe stimulus processing as a function only of ‘bottom-up’ processes and, 3) to me, it seems to view preconscious neurological processing as hard-wired and not under conscious control. In response to the first two points, I argued earlier that ‘self’ can be viewed as teleological process conducted by neurobiological mechanisms. I assert that the same is true of other ‘higher level cognitive events’, including the sense of agency. If these ‘events’ are ontologically reducible to mechanistic neurobiological models (Carver, 2006) I suggest that, by nature of their multi-level activity, it therefore follows that they should not be construed in simplistic ‘bottom-up’ and ‘top-down’ terms.
However, the third point about control is more challenging. Wegner (2002) advances a similar view in his theory of Apparent Mental Causation. Wegner posits that conscious will can be defined as an illusion generated by causal inference between conscious thought prior to an action and self-awareness of enacting the action. Wegner argues that interaction between cognitive and self-authorship neurological mechanisms generates 'illusory agency' – i.e.

"Experiences of conscious will thus arise from processes whereby the mind interprets itself - not from processes whereby mind creates action. Conscious will, in this view, is an indication that we think we have caused an action, not a revelation of the causal sequence by which the action was produced" (Wegner, 2004, pg649 - original emphasis).

Wegner's suggestion of illusory will is critiqued by Nahmias (2005) who argues that, amongst other criticisms, this view of agency does not easily account for agentic capabilities like 'distal intention' (i.e. planning ahead) and is, therefore, too simplistic. However, the most significant evidence in support of causative agentic control comes from research demonstrating deliberate conscious modulation of subconscious processing. These studies provide evidence to suggest that neuronal mechanisms operating at subpersonal level very much operate within one's awareness and control, which is not in keeping with a Bandurian view.

Neuroplasticity and neurogenesis are well-documented (e.g. Pascual-Leone et al., 2005) and, to me, it come as little surprise that learning a new skill

---

Possibly using Frith's (1992) proprioceptive system outlined previously, although Wegner also suggests other somatic-feedback systems too (see Wegner, Sparrow & Lea, 2004).
(e.g. violin playing [Elbert et al., 1995] or juggling [Draganski et al., 2004]) involves neurological growth and cortical remodelling (Ilg et al., 2008). Critically, however, it has been reported recently that the same processes may also be controlled consciously - i.e:

"With appropriate training and effort, people can systematically alter neural circuitry associated with a variety of mental and physical states" (Schwartz, Stapp & Beauregard, 2005, pg2).

This self-directed neuroplasticity has been documented in a number of clinical conditions such as OCD (Schwartz et al., 1996; Schwartz 1998), depression (Brody et al., 2001; Martin et al, 2001), Wernicke's aphasia (Musso et al., 1999) and spider-phobia research (Paquette et al., 2003). In these studies the positive impact of mental training and coping exercises was associated with significant changes cortical fMRI activity pre- and post-training, implying that "people deliberately and consciously alter[ed] neural circuitry" (Schwartz & Begley, 2002, pg11). More evidence supporting ‘thinking induced plasticity’ (Schwartz & Begley, 2002) can be found in non-pathological studies. For example, Hölzel and her associates (Hölzel et al., 2008; 2011) observed that regular periods of meditation led to increases in grey matter density and cortical thickening; whilst others (De La Fuente-Fernandez et al., 2001) have published evidence to suggest that the placebo effect may be attributable to self-induced dopamine release. Numerous studies (e.g. Ochsner et al., 2002; Beauregard, 2001; see also Beauregard, 2007) have reported that conscious influence on emotional state is linked with neurological-based changes in affective regulation (see, for example, the sections in the previous Chapter on reappraisal processes). Therefore, in contrast to
Bandura’s (2011) assertion, I posit that such subconscious processes are not only exceptionally malleable, but that they may also be subject to significant direct and indirect influences from the activities of other entities within the agentic neurobiological mechanism.

Thus, I conjecture that the neurobiological agent is anything but an automaton: rather, through a system analogous to ‘second-order control’ (Bandura, 2006a; 2006b: 2008), I have argued that the agentic neurobiological process may consciously self-guide by indirectly modulating subconscious entities to achieve a full measure of control over thoughts, feelings and emotions. Therefore, although my views on the generation of agency are very different to Bandura’s, we arrive at precisely the same conclusion: as humans we “make a causal contribution to our own motivation and action” (Bandura, 1989, pg1175) via our “beliefs, values and goals, which set up a meaning system within which we define ourselves (Dweck, 2000, pg138-139). Thus, in the words of Beauregard:

“With the emergence of self-consciousness, self-agency, and self-regulatory... capacities, evolution has enabled humans to consciously and voluntarily shape the functioning of their brains. These advanced capacities allow humans to be driven not only by survival and reproduction but also by complex sets of insights, goals, and beliefs” (Beauregard, 2007, pg222).

Efficacy beliefs

Human agency is believed to centre around four such beliefs, or ‘features’, which “define what it means to be a human” (Bandura, 2001b, pg6). These features are intentionality, forethought, self-reactiveness
and self-reflectiveness (Bandura, 1997; 2008). Each of these core aspects of agency are known to be affected by sociostructural and environmental influences (Bandura, 1986b), which interact with internal personality factors and behaviour to form the deterministic reciprocal triad that Bandura (1997) defines as underpinning human agency.

Although a lot of evidence points to significant contributions from behaviour and the environment in the generation of agency (Bandura, 1986b), Bandura has consistently asserted that personality factors and, specifically, “beliefs of personal efficacy constitute the key factor of human agency” (Bandura, 1997, pg3).

Bandura defines Self-Efficacy (SE) as:

“People’s beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives” (Bandura, 1994, pg71).

In other words it is the measure of confidence or conviction “that one can successfully execute the behaviour required to produce [a desired] outcome” (Bandura, 1977, pg193). Such beliefs should not be confused with the component repertoire skills they encompass, rather:

“Self-efficacy is concerned not with the number of skills you have, but with what you believe you can do with what you have under a variety of circumstances” (Bandura, 1997, pg37).

For example, Bandura (1984) draws the example of driving a car. He argues that a person’s belief in their ability to drive, whilst based in part on their ability to steer, change gears, interpret road signs etc,
comprises considerably more than the sum of the subskill components. Thus, in this example, driving SE constitutes a meta-belief in its own right—“investigators would do well to follow the dictum that the whole is greater than the sum of its parts” (Bandura, 1997, pg37). In keeping with the same driving analogy it is important to acknowledge that SE beliefs are believed to be very task- and context-specific (i.e. they do not generalise well from one domain to another), thus whilst a person may have high SE driving a car through the country at the weekend, they may be very reluctant to take on the motorway in the rain at midnight.

Although SE beliefs do not appear to vary greatly in generality, they have been found to differ considerably in level, which is a measure of perceived task demand (Bandura, 1997) or, in simple terms, how easy the task is perceived to be. Thus, as Chase (2001) observed, only individuals with high SE choose to undertake tasks that they perceive challenging (Chase specifically cites physical activities like rope-climbing). In similar vein, evidence suggests that SE beliefs also vary in strength, which describes the tenacity of SE to overcome disconfirming experiences. Although separate constructs, SE strength overlaps considerably with persistence (Bandura, 1997), behavioural choice (Bandura, 1977) and intention (Wulfert & Wan, 1995), which are all hypothesised to act synergistically with SE to affect performance.

SE beliefs have been reliably reported both to predict attainment (Bandura, 1997) via the causative triadic interactions outlined above (Bandura, 1986b). In this way SE is viewed as deterministic, rather than prognostic: “self-efficacy beliefs determine how people feel, think, motivate themselves and behave (Bandura, 1994, pg2 - original emphasis).
This effect is well evidenced (Bandura, 1992). For example, Baron (1988) conducted a prospective study into the effect of different types of feedback on undergraduates’ performance. He found that destructive, untargeted feedback led not only to decreased SE on further tasks, but also to a significant reduction in the quality of work the students produced in later assignments. Meta-analyses (e.g. Guzzo et al., 1985; Kluver & De Nisi, 1996) also support Baron’s observation that feedback intervention moderates performance indirectly via increased SE (although see Kluver & De Nisi, 1996 for other suggestions).

Bandura (1977, 1986b) posits that the effect of SE on performance occurs via mediation between knowledge and action. Apropos, it has been hypothesised that SE affects performance in a large number of ways, which include; cognitive appraisal of personal, environmental, and behavioural variables (Bandura, 1977); changes in performance goal-setting and intra-performance coping strategy (Bandura, 1977; 1986b; 1992), discarding faulty self-regulatory strategies more quickly (Zimmerman, 2002), increasing persistence (Bandura, 1991a), motivation (Bandura, 1992) and emotional and physiological arousal (Bandura, 1986b; 1991; 1992), or working more accurately than those with lower efficacy (Bandura, 1986b, 1992). A full summary of the effect of SE on performance is presented in Appendix A.

More significant than the effect of SE on performance is the weight of literature supporting a link between low SE and ill health. For example, in a prospective study McFarlane and colleagues found that people who scored lower on measures of self-efficacy showed substantially more symptoms of depression in later life (McFarlane et al., 1995; see also
Petersen et al., 1993). Similar results were found by Cowen (Cowen et al., 1991), who also showed that low SE scores are strongly correlated with negative social-emotional and behavioural functioning. Their results have been replicated both in retrospective correlation analyses (Ehrenberg, Cox, & Koopman, 1991) and prospective studies (Bandura et al., 1999), including experiments in which negative affect traits are controlled for (Muris, 2002). Other studies have found correlations between low SE and mental illnesses such as anxiety disorders (Matsuo & Arai, 1998; Yue, 1996) and addiction, including alcoholism (Taylor, 2000) gambling (Casey et al., 2008) and drug dependence (Stephens et al., 1993). However, the link between low SE and the development of clinical conditions is not limited solely to mental illness, as SE has been implicated in the regulation of the specific immune response (O’Leary & Brown, 1995; Caserta et al., 2011; see also Bandura, 1997), the development of cardiovascular disorders (Cutrona & Troutman, 1986) and has recently been reported to predict heart failure in people suffering from CHD (Sarkar et al., 2009).

Similarly, considerable evidence implicates SE in the development of both acute and chronic stress. For example, self-reported SE is repeatedly found to correlate negatively with self-reported stress levels. This has been reported in studies involving university students (Hackett et al., 1992; Solberg et al., 1993; Gigliotti & Huff, 1995; Solberg & Villarreal, 1997; Torres & Solberg, 2001), teachers (Cadiz, 1989; Betoret, 2006; Vaezi & Fallah, 2011), nurses (Schaubroecck & Merritt 1997; Lo 2002), mothers of young children (Teti & Gelfand, 1991), carers (Pinquart & Sörensen, 2003) and across a wide variety of occupational and professional environments as well (Jex & Gudanowski, 1992; Matsui &
Onglatco, 1997; Jex & Bliese, 1999; Prati et al., 2010). However, such correlations, by themselves, do not demonstrate a causal link between SE and stress. Indeed, on these grounds, Lee (1989) has criticised the SE construct in three regards: 1) SE is vague, which effectively makes it an unfalsifiable concept\(^{50}\), 2) SE is commonly used as a behavioural descriptive model (rather than an explanatory theory), and 3) SE is an a priori hypothetical construct invoked to help understand a pattern of observed behaviour and, therefore, cannot be an independent variable in itself. Whilst Lee’s first point is well made (although see Bandura’s [2012] comment about variation in instruments designed to measure SE), Bandura (1982; 1997) has repeatedly asserted, contrary to Lee’s criticism, that SE does influence stress causally. In his most recent book, Bandura (1997) cites numerous studies in which efficacy beliefs are purported to have a causal effect on behaviour via artificially raising or lowering SE through false positive- or negative feedback (e.g. Bouffard-Bouchard, 1990; see also Appendix A). In stress research concomitant procedures have been observed to have a similar effect on stress: for example, Holroyd and colleagues (Holroyd et al., 1984) found that manipulating SE feedback during an EMG biofeedback training programme successfully reduced tension-induced headaches, even when the positive feedback was spurious. Similar findings in other false-feedback experiments (Neufeld & Thomas, 1977; Litt, 1988; Kores et al., 1985; Manning & Wright, 1983) also support a causal role for SE in limiting the development of stress. This conclusion is further reinforced through the success of stress-reduction programmes, where students undertaking a curriculum designed to boost SE have reported significant reductions in stress. This is particularly well established in outdoor training schemes (Thurber et al.,

\(^{50}\) This also explains the ease of associating SE with behaviours in correlation studies.
2007; McKenzie, 2000; Masten & Coatsworth, 1998; although see Sheard & Golby [2006], who present an alternate explanation of the success of outdoor educational programmes), and also in recent classroom-based studies, which have reported similar findings (e.g. Lohaus, 2011; Hampel et al., 2008; Frydenberg, 2004; see also Kraag et al., 2006), prompting the conclusion that SE training not only improves academic performance, but may also have a causal effect on stress reduction (Mori & Uchida, 2009). Therefore, as Bandura posits;

"People who regard themselves as highly efficacious act, think, and feel differently from those who perceive themselves as ineffectual" (Bandura, 1997, pg395).

In this way, SE is believed to exert a causal, protective role in limiting the development of stress when people are exposed to stressors. The key, question, therefore, is how might this happen?

Most research to date attempts to explain this 'protective' nature of SE in limiting stress development in terms of adaptive changes in coping style (e.g. Taylor, 1989; Rothbaum et al., 1982; Brandtstädtter, 1992; Lazarus & Folkman, 1984; Lazarus, 1999; see also Suls & Fletcher, 1985; Parker & Endler, 1996; Skinner et al., 2003; Snyder, 1999; Lazarus, 1993b). Although robust evidence exists to link the adoption of adaptive coping style with reduced stress (e.g. Mikolajczak et al., 2007) and in attenuating stress-related illness\(^5\) (e.g. Temoshok et al., 2008; see also Penley et al., 2002), a wealth of studies have repeatedly demonstrated

\(^5\) It should also be noted that coping-based intervention programs have had some success in limiting stress (e.g. Antoni et al., 2006; Fife et al., 2008; see also de Riddler & Schreurs, 2001).
that coping strategy is heavily and independently influenced by antecedent appraisal processes (Bjorck et al., 2001; Portello & Long, 2001; Pakenham, 2001; Chung et al., 2001; Mikulincer & Victor, 1995; Anshel & Wells, 2000) and that, by itself, coping style is not predictive of the development of stress (Chung et al., 2001), or of stress-related disease (Blalock & Joiner, 2000). Thus it appears that, whilst SE-induced adaptive coping is likely to play a role in reducing stress, its effect in stress genesis may be more limited. Considering this supposition in conjunction with other, more general criticisms of coping theory in stress research (e.g. Perlin, 1991; Snyder, 1999), I posit that SE may not exhibit its protective nature exclusively via changes in adaptive coping; instead I propose that the protection may also occur through changes in the appraisal of stressors.

In the previous Chapter I explained that dual process models of stress identify two discrete and functionally independent stressor appraisal processes: those that are automatic and those that are conscious. Thus far research appears to have focused predominantly on conscious stressor appraisals. Indeed, many studies have reported that the negative correlation between SE and stress is clearly associated with a change in conscious stressor appraisal (Chemers et al., 2001; Lazarus & Folkman, 1984; Pintrich & De Groot, 1990), which supports my conjecture that SE might exhibit at least part of its protective nature via changes in

---

52 Their work focused heavily on depression. Little evidence exists for other stress-related diseases.
53 See also Coyne and Gottlieb (1996), Coyne (1997) and Somerfield (1997), who criticise the methodologies employed in coping research. Others draw attention to the inability of coping paradigms to explain preconscious reactions to stressors (Somerfield & McCrae, 2000) and the repeated failure of coping research to inform, or to engender positive change in clinical practice (Coyne & Racioppo, 2000).
the way in which stressor stimuli are evaluated. However, I have been unable to identify any research investigating a link between SE and preconscious appraisals processes. Thus, as explained in Chapter 2, not only does a 'gap' in stress research exist in establishing whether stress regulation includes change to preconscious appraisal mechanisms, but an additional 'gap' exists in determining the involvement of SE on such preconscious appraisals. Therefore, the role of this study is to take the first few steps to explore these 'gaps' in an attempt both to expand our understanding of the protective nature of SE in limiting stress, but also to draw the attention (and resources) to an area of research that could potentially help to achieve my over-arching goal of witnessing the establishment of a national curriculum that incorporates a focus on stress resistance, or anti-stress training.

Summary and research questions

By citing work in the field of self-directed neuroplasticity, I argued earlier that, as efficacious agents, humans are capable of consciously and deliberately altering the neurological processes that underpin cognition. By drawing together evidence in the field of emotion regulation I further asserted that affect may be altered in similar ways - i.e. the conscious and preconscious appraisal processes that generate emotion may be regulated significantly to manage resultant emotions. However, in contrast to the self-directed neuroplasticity paradigm, the functional manipulations that regulate emotion occur not only consciously, but also preconsciously, without our awareness. As I have suggested that stress shares functional equivalence with affect regulation (and provided significant evidence to support this position), I further advance that the
emotional regulation processes outlined here also apply to the management of stress. If one accepts the hypothesis presented - that stress is an emotion and that agentic beliefs can alter cognitive processes - I present the logical sequitur that SE may manipulate stress via an effect on stressor appraisals. I further assert that, whilst SE may interact with processes that regulate conscious appraisals, a second significant stress-limiting effect could also occur through unconscious regulation of preconscious stressor appraisal processes. Therefore, I propose that SE may also provide some of its protective effect by unconsciously ‘toning down’ the preconscious appraisal of stressor stimuli. In other words, people with high SE may simply not perceive stimuli to be stressors, which is in keeping with the conclusions of Bandura (1997), Lazarus (1999) and others (e.g. Lazarus & Folkman, 1984), albeit for very different reasons!

To explore my assertions further, I propose the following research questions, which will form the focus of this study:

1. To what extent does academic SE limit the development of acute academic stress?

2. To what extent does acute academic stress exposure affect the preconscious appraisal of academic stressor stimuli?

3. To what extent does academic SE affect the preconscious appraisal of acute academic stressor stimuli?
Choosing the research methodology

Post-positivist methodologies seek to reveal causal descriptions (Shadish et al., 2002) between manipulated-, controlled- or empirically observed variables, most frequently through the use of objective, randomised experimental procedures (Guba & Lincoln, 1994). Such protocols are:

"Widely considered the gold standard because they are expected to produce an estimate of the mean treatment effect on a given dependent variable that deviates from the true value only by random error, which is kept small when statistical power is adequate" (Wilson & Lipsey, 2001, pg413).

However, consistent with the post-positive notion of multiple operationalism (Campbell & Fiske, 1959):

"Although randomised experiments are often seen as the golden standard... there are many situations where the use of experimental design is not suitable or [is] simply impossible" (Aussems et al., 2011, pg21).

Thus the researcher is compelled to draw from other post-positivist methodologies, such as modified experimental protocols (Cook & Campbell, 1979), or other quantitative comparative procedures (e.g. Cochran, 1965).

As an example, consider the clinical trials process for testing new pharmacological agents. This procedure entirely typifies Shadish's 'true' experimental methodology (Shadish et al., 2002) as it employs a system of pre- and post- intervention testing, the use of control groups and,
crucially, the randomisation of participants and of treatment via placebo and double-blind protocols. However, this well-established process becomes entirely inappropriate if, for example, the efficacy of the drug being tested has already been partially established. Under such circumstances ethical considerations would prevent the researcher from withholding the treatment from members of a control group, which immediately removes the possibility of the trial following the standard post-positivist clinical trial protocol (Harris et al., 2006). In similar fashion, access to a limited number of participants, or the inability to randomise those participants into control and experimental groups undermines the feasibility of following an experimental methodology, and thus necessitates the use of methodologies other than randomised experiments (Cook & Campbell, 1979). In such circumstances, quasi-experimental methodologies are frequently used (Cook & Campbell, 1979) as they allow causal hypotheses to be tested;

"In circumstances where full control over the scheduling of experimental stimuli (the when and to whom of exposure and the ability to randomise exposures) which make a true experiment [are not] possible" (Campbell & Stanley, cited in Shadish et al., 2002, pg14).

In this research all three of the scenarios illustrated above apply. Firstly, by definition, studies conducted within one school are subject to the limitations of a small sample size (at least on the kind of scale used in large randomised experimental studies). Secondly, as ethical guidelines insist that “participants must be protected from stress by all appropriate measures” (British Psychological Society, 2012, pg1), this removes any possibility of establishing an independent stressor stimulus of the kind
necessary for a ‘true’ experimental protocol. Lastly, as a consequence of the previous limitation, the procedure for allocating participants to experimental and control groups cannot be operationalised arbitrarily: in other words, it was not possible to follow the randomised grouping protocol that lies at the heart of the ‘true’ experiment (Guba & Lincoln, 1994). In the words of William Shadish:

"Those of us who toil in the trenches of fields like psychology, education and economics know that random assignment is what we would like to do, but that quasi-experiments are what we are sometimes forced to do for practical or ethical reasons" (Shadish & Cook, 1999, pg294).

Therefore, in summary, although my ontological preference would be to uphold the ‘golden standard’ of post-positivist research (Shadish & Ragsdale, 1996) – the randomised experiment – for the reasons outlined above I chose to adopt the next best epistemic approach by following a quasi-experimental methodology in this research.

Although the validity of quasi-experiments has been questioned (Glazerman et al., 2003) on the grounds that they offer limited scope to draw causal inference between intervention and outcome, a large body of research exists to support their role as a valid methodology for use in social scientific research. For example, in a pioneering large-scale study Dehejia and Wahba (1999) modelled the effects of a labour training programme on post-intervention earnings using both a randomised procedure and non-randomised trials. Their work (and its subsequent ratification by others e.g. Hill et al., 2004) provided strong supportive evidence that quasi-experimental and randomised studies can develop
similar effect sizes and equivalent predictive power. Although the conclusions of these studies have been debated (e.g. Shadish et al., 2008), a number of authors have replicated their findings using propensity scoring paradigms (e.g. see Shadish et al., 2006) and multiple regression analyses (Steiner et al., 2010), which have been found either to reduce non-randomisation bias significantly (Shadish et al., 2008), or to remove it entirely (Aiken et al., 1998; Black et al., 2007). Other evidence supporting the equivalence of randomised experimental and quasi-experimental methodologies can be drawn from computer simulations (Drake, 1993) and doubly-randomised preference trials (Rücker, 1989), which Janevic and colleagues (2003) also found to reveal no significant difference between randomised- and non-randomised trial effects.

However, whilst the studies mentioned above provide peripheral support, the most substantive evidence corroborating the validity of quasi-experimental methodology comes from meta-analytic studies. For example, in a simple comparison of outcomes, Lipsey and Wilson (1993) found little difference in treatment effect between quasi-experimental and true experimental studies: a result later confirmed by Heinsman and Shadish (1996). Whilst this research has been criticised for its summative nature (i.e. it compiled outcomes only and made no attempt to regress confounding covariates – Shadish et al., 2008) other, more statistically rigorous projects, have taken these limitations into consideration. For example, in a meta-analysis of 47 quasi-experiments and 52 randomized experiments Heinsman (1993, cited in Shadish & Heinsman, 1997) found that randomised experiments yielded only 'slightly higher' effect sizes, whereas other studies have found no significant
differences at all (Wampold et al., 1997; Stiles et al., 1986; Meldrum, 1998; although see Shadish & Ragsdale, 1996; Glazerman et al., 2003, who did find significant differences between methodologies). This effect is further confirmed by Wilson and Lipsey (2001), who conducted the largest meta-meta-analysis to date (of 319 meta-analyses) and found:

"Virtually no difference, on average, between the results from nonrandomized comparison group designs and those from randomised designs" (Wilson & Lipsey, 2001, pg424).

That said, their study also revealed large discrepancies between the validity of some quasi-experimental studies, leading them to urge caution in assuming congruence between quasi-experimental and randomised experimental methodology:

"These results underscore the difficulty of detecting treatment outcomes [and] the importance of cautiously interpreting findings from a single study" (Wilson & Lipsey, 2001, pg413 - abstract).

This is particularly the case given the influence of unidentified or uncontrolled covariates (Steyer et al., 2000; Shadish & Ragsdale, 199654) and the limitations of statistical regression (Shadish et al., 2002). Therefore, whilst quasi-experiments do, under carefully planned circumstances, have equivalent validity to randomised experiments, my choice to follow this methodology comes with the strong proviso that I plan the study as carefully as possible and take every precaution possible.

---

54 Their study found that variation could be reduced by half if particular covariates were taken into account.
to limit the influence of confounding variables generated through the use of non-randomised groups.

Choosing the quasi-experimental design

Quasi-experiments establish causal descriptions (Aussems et al., 2011) using protocols:

"That have treatments, outcome measures and experimental units [participants], but do not use random assignment to create the comparisons from which treatment-caused change is inferred" (Stouffer & Campbell, cited in Cook & Campbell, 1979, pg6).

In this research the availability of modular A-level examinations in the January 2013 examination season determined the assignment of participants into control and experimental groups which, although a number of different quasi-experimental procedures exist (Shadish et al., 2002), necessitated the choice of a non-equivalent group design for the methodology. Of these, the untreated control group design with one pre-test was selected (see Figure 1 below) for the reasons given below.

![Figure 1: Diagram representing the untreated control group design with one pre-test quasi-experimental methodology (adapted from a similar diagram in Shadish et al., 2002).](image-url)
This methodology has two distinct advantages over alternatives. Firstly, the inclusion of a pre-test procedure is widely recognised to improve quasi-experimental design (Shadish et al., 2002), as it enables the influence of ‘noise’ from pre-treatment variables to be removed from the measurement of the dependent variable, either by use of ANCOVA or statistical regression analyses. The second advantage is that this methodology makes use of a control group. Control groups literally provide control over the independent variable, which “helps separate the effects attributable to a treatment from the effects attributable to irrelevancies that are correlated with a treatment” (Cook & Campbell, 1979, pg31). This would not be possible, for example, following a one-group pre-test / post-test design. Similarly, control groups also remove the possibility of regression towards the mean, which is a common problem for quasi-experimental methodologies that use a single group design (Shadish et al., 2002).

Therefore, in summary, quasi-experimental methodologies that use both control groups and pre-testing are recognised as being of ‘sounder design’ (Harris et al., 2006) than most other quasi-experimental designs and, as such, are the preferred methodology for use in this study. That said, Cook and Campbell (1979) identify other methodologies with the potential to develop more statistical power than the methodology employed here. However, such protocols frequently necessitate the inclusion of switching replications (Shadish et al., 2002) within the methodology. This involves utilising a second control group who, in this scenario, would have been required to sit the public examinations at a different time in the protocol to the other groups. Whilst the Joint Council for Qualifications (JCQ) do permit some flexibility in the scheduling of public examinations (see JCQ,
2012), this strictly applies to situations in which students have two or more public examinations scheduled at the same time on the same day, or when more than six hours of examinations are scheduled in the same day. Other changes to examination schedules are not permitted. Clearly, then, it is neither feasible, nor ethical to move the dates of public examinations to suit a research methodology, which explains why protocols with switching replications, despite their advantages, were not incorporated into this study.

In similar vein, methodologies incorporating repeated-measures offer scope to develop the internal validity of a study because they allow for repeated testing of the same participants under identical conditions - a protocol very similar to test-retest procedures. Effectively, such methodologies control for variations between participants and would thus offer considerable potential to develop the generality of the conclusions of this study. However, as with methodologies employing switching replications, pragmatic considerations prohibited the use of such measures in this study. For example, many participants in the study sat just a single public examination in the January 2013 season, which precluded the option of repeated stressor exposure. Even for students who took multiple exams, the short intervals between papers (often only a few hours) made repeated testing phases both impractical and, arguably, unethical. One potential alternative would have been to arrange for some participants to take an additional raft of mock examinations, either a few weeks after, or a few weeks before, their public exams. I rejected this option for three reasons; a) on the grounds of the significant disruption to the teaching timetable such additional exams would inevitably cause, b) ethical reasons associated with the previous point, particularly as the
rationale for such an intervention bares little direct educational benefit to the pupils and, c) because of the implicit difference in stimulus valence between a ‘real’ public exam and a ‘fake’ mock exam - i.e. because an additional mock, however similar to a public exam, is fundamentally not a public exam, it cannot constitute an identical replication of stressor exposure and is, therefore, self-defeating in a repeated-measures methodology. Thus, for the purposes of this study, I limited testing to just a single phase of examinations.

A final reservation with the choice of methodology is the use of a single pre-test. Campbell advocates the use of multiple pre-tests (Cook and Campbell, 1979, pg37) because they "assist in controlling time-varying confound effects," such as threats associated with differential rates of participant maturation (Shadish et al., 2002), or the identification of regression artefacts generated by analysing only part of a natural cycle (Cook & Campbell, 1979). In both cases quantitative assessment of time-dependent 'background' effect is possible, which can then be eliminated from the overall effect statistically (West et al., 2000). In a similar way, multiple pre-tests allow for assessment of instrument internal reliability (commonly by calculation of Cronbach's alpha, or test-retest procedures) which, as pre-test error is known to combine synergistically with error generated through non-random grouping, can be statistically factored out of post-treatment assessment to avoid 'flattening' of regression lines. As the flattening effect is known to distort accurate assessment of the influence of the independent variable (particularly in methodologies using only a single pre-test - Trochim, 2006), the clear conclusion is that quasi-

---

55 This point is arguably amplified by conducting the study in an independent school, where the parents pay fees for the provision of their daughter's education.
experimental methodologies employing multiple pre-tests are advantageous over those using a single phase of data collection. This obviously begs the question, 'why were multiple pre-tests not used in this research?'

The simple answer is that I had intended originally to follow an untreated control group design with two pre-tests: my 'plan' scheduled the project to run during the January 2014 examination season, providing ample time to recruit participants, secure necessary consents, and to run both pre-test phases of data collection in the last months of 2013. However, on the 9th November 2012 the Office of Qualifications and Examinations Regulation (Ofqual) suddenly and unexpectedly announced that, as part of on-going A-level reforms, "from next September, students will only be able to sit AS and A level exams in the summer" (Ofqual, 2012), which meant that the last available January modules would occur in the January 2013 examination season. With plans thoroughly scuppered, this unwelcome announcement left me with four viable options: a) move the study to a summer examination season, b) switch the study to an internal examination season, c) recruit younger participants, or d) bring the project forwards a year.

As virtually every student in every school in the country sits public examinations in June, following the first option would have meant abandoning the use of a control group. Whilst other quasi-experimental methodologies exist that adopt a single group format, the clear guidance is that, on the relative hierarchy of quasi-experimental design (Harris et al., 2006), methodologies using control groups develop more validity (Cook & Campbell, 1979) than those that do not. This effectively precluded
following the first option. A similar reservation applied to the second option, with the added limitation that mock examinations (or internal examinations) are intrinsically less stressful than the real thing. Therefore, I dismissed the option of designing the methodology around internal exams. I considered carefully the option of recruiting participants from Year 10, as some exam boards offer the option of sitting early modules in some GCSE courses (e.g. Edexcel Geography, OCR 360 Science or AQA Drama). Whilst this arrangement would provide the option of a Control group (most Year 10 students do not sit public examinations in this year), I identified three significant limitations with this option, 1) the development of SE is age-dependent (see later section on participants) and is less defined in younger people, 2) additional ethical implications would need to be considered, were younger participants to be recruited to the study, and 3) the small number of students sitting public examinations in Year 10 would necessitate the recruitment of pupils from a number of schools to reach a sufficient sample size to generate the guideline 80% statistical power (Lipsey & Hurley, 2009 - see later). Taken together, these reservations led me to dismiss the third option as well. This effectively left moving the study forwards by a year as the only viable, practical solution. The caveat to this decision was that, given the time constraints of recruiting participants ethically, I was unable to schedule two phases of pre-testing and, therefore, the study followed an untreated control group design with just a single pre-test.

Therefore, in summary, I posit that the choice of an untreated control group design with one pre-test constituted the most valid quasi-experimental methodology to employ in the context of the research, both for pragmatic reasons and because it was the highest ranking available
option on the relative hierarchy of quasi-experimental designs (Harris et al., 2006).

**Application of the quasi-experimental design**

Participants were allocated non-randomly into two non-equivalent groups, which I refer to from this point forwards as the *Experimental* group and the *Control* group. Allocation was determined by the availability of modular AS and A2 A-level examinations in the January 2013 examination season. Students with examinations (N=70) constituted the Experimental group and students with no examinations (N=44) constituted the Control group. As examinations are widely recognised to be the most significant stressor adolescents experience at school (Burnett & Fanshawe, 1997; Kohn & Frazer, 1986), the premise behind the grouping procedure was that the Experimental group would experience acute academic stress as a consequence of exposure to academic stressor stimuli (represented by 'intervention X' in Figure 1). By contrast, participants in the Control group were not exposed to academic stressors and, therefore, did not develop acute academic stress.

As summarised in Figure 2 below, members of both groups participated in two discrete and identical phases of data collection, which I refer to as pre-stress (PrS) testing and post-stress (PoS) testing. PrS testing occurred 6-7 weeks before stressor exposure, whilst the PoS phase took place in the week of the examinations themselves. In keeping with guidance from Shadish and colleagues (Shadish et al., 2002) participants undertook PrS and PoS testing at approximately the same time of day, by completing instruments over two lunchtime sessions (either 12:40 - 13:00
or 13:00 – 13:20), which were available every day during the week of testing. Participants were instructed to come for testing on a specific day, but were free to choose which lunchtime session they came to. This procedure intended to limit clashes with other lunch commitments and, thereby, to increase participant retention.

PrS and PoS testing took place in one of the school IT suites in groups of 5-20 participants. Each computer workstation was screened from the others using free-standing A3 display boards, which stopped participants from talking to each other, or from distracting each other during testing. It also served to preserve participant privacy.

Control Group:

Experimental Group:

Figure 2: Diagram representing the three phases of data collection.
Each phase of data collection consisted of three different instruments: the Examination Self-Efficacy Scale (ESES), the Perceived Stress Scale (PSS10) and the Stressor-Appraisal Implicit Association Test (SA-IAT). The instruments were completed over a period of 10 - 20 minutes in the order shown in Figure 2. The PSS10 and ESES were completed on paper (see Appendices E & D respectively), whereas the SA-IAT was hosted on the school's VLE and, accordingly, was completed using a networked PC. Although I document the structure, function and validity of these instruments separately in this Chapter, the primary function of these instruments is as follows:

SA-IAT: measures the preconscious appraisal of stressor stimuli
ESES: measures academic SE
PSS10: measures perceived stress

Relating the design to the research questions

1. To what extent does academic SE limit the development of acute academic stress? The impact of academic stressor exposure on stress can be measured by performing an ANCOVA analysis between Control and Experimental groups with PoS stress levels as the dependent variable and PrS stress level as a covariate. Any contribution of academic SE towards the development of stress can be measured using a Pearson's Product Moment Correlation Coefficient for each participant group, which assess the degree to which academic SE is correlated with PoS stress. Lastly, any differences revealed by the regressions can be assessed for significance through application of a Fisher's r-to-z transformation calculation.
2. To what extent does acute academic stress exposure affect the preconscious appraisal of academic stressor stimuli? Similar to the previous question, the difference in post-stressor exposure measures of implicit stressor appraisal can be contrasted between Control and Experimental groups by conducting an ANCOVA analysis on PoS SA-IAT scores, with PrS SA-IAT data as a covariate.

3. To what extent does academic SE affect the preconscious appraisal of acute academic stressor stimuli? The correlation between SE and implicit appraisal can be contrasted between Control and Experimental groups in both pre- and post-stressor exposure phases using Fisher’s r-to-z transformation calculations. For any significant differences, the contribution of academic SE towards implicit stressor appraisal can be assessed using a linear regression analysis.

**Considerations and Control Procedures**

**Participants.** In order to gauge the approximate number of participants required for this study I conducted an analysis of statistical power using G*Power 3.1 (Faul et al., 2009). Using Cohen’s (1977) recommended values of Type I and Type II error (respectively, α = .05 and β = 0.2), the minimum number of participants required to achieve 80% power (the conventional ‘minimum target’ of power for statistical analyses – Lipsey & Hurley, 2009) in an ANCOVA analysis was calculated as between 52 and 128 per group. This assumes a large to medium effect size (i.e. θ lies within the range .25 ≤ θ ≤ .4), as defined by Cohen (1992). These calculations are displayed graphically in Figure 3.
All schools develop a unique personality (Schein, 1985), or culture - “the way we do things around here” (Bower, 1966, pg12) - which is “manifested in people’s patterns of behaviour, mental maps, and social norms” (Paterson & Deal, 2009, pg18). This school ‘culture’ varies significantly from school to school (Paterson & Deal, 2009) - even when schools with students from broadly similar socioeconomic backgrounds are contrasted (Sirin, 2005) - and has been shown to exert a significant contributing effect on students’ SE (Artelt et al., 2003). This effect is independent of other factors known to affect average student SE, including socioeconomic status (McConney & Perry, 2010) and the academic standing of the school (Chiu & Xihua, 2008). Taken in conjunction with research from other contexts demonstrating the influence of nationality (Kim & Park, 2006) and culture (Oettingen & Zosuls, 2006) on SE, I reasoned that the 52 - 128 participants should be recruited from the same school and should, therefore, share a similar sense of school

Figure 3: Graphs representing G*Power power analysis for minimum sample size (left, $\theta = 0.4$, right, $\theta = 0.25$)
culture. This decision to recruit all participants from the same school was made in part for pragmatic reasons, in part for ethical reasons (mostly associated with access to other schools), but mostly as an attempt to limit the confounding effect of school culture on SE. Therefore, I chose to limit the study to the school in which I worked. This decision had a number of additional benefits:

a) As the school selected is an independent and academically selective fee-paying day school, the influence of many of the other social factors known to affect SE (e.g. socioeconomic status, comprehension of English or General Intelligence [GI]) was limited as far as was possible.

b) Gender is well-known to affect the development both of academic SE (Britner & Pajares, 2001; Pajares & Valiante, 2001; see also Hansen, 2009; Pajares, 2002) and of acute- (Matud, 2004) and chronic stress (Cleary, 1987), as well as playing a significant role in the pathogenesis of stress-related illness (Weissman & Klerman, 1987). As the school selected is a single-sex girls’ school, the confounding influence of gender was effectively controlled. However, this introduced the caveat of the effect of hormone cycles on self-reported stress and SE. However, as this effect was dispersed evenly across both participant groups (Control and Experimental group) and, as ANCOVA assesses variation between groups, the influence of hormone cycles was, therefore, eliminated statistically by the use of ANCOVA.

c) Similar to the previous point, age has been shown to exhibit a marked effect on the development of academic self-efficacy (Anderman et al., 1999; Urdan & Midgley, 2003) and stress (Rudolph & Hammen, 2003;
Rutter, 1981). Therefore, in order to reduce the effect of these variables, I chose to limit participant recruitment to within a single key stage. For ethical reasons, the availability of a large number of public examinations in January 2013 and because the size of the Sixth Form in the school selected (150 girls) exceeds the upper figure in the desired sample size (128), I elected to limit the study to Year 12 and 13 pupils at the chosen school.

**Ethics.** Prior to collecting data, I sought approval for the design of the study from the Ethical Research Council at Brunel University. Approval was granted and a copy of this letter is included in Appendix C.

Based on Creswell's (2003) guidance, before beginning any data collection all participants were given a full briefing (see Appendix D), which explained their right to withdraw from the study at any point, their right to withdraw their data from the project, if they so desired, their right to ask questions and to obtain a copy of the results and an outline of the benefit of the study. The participants were then asked to sign a consent form (see Appendix E), which indicated that they understand the procedure and gave their informed consent to take part. As the majority of participants were under the age of 18, I requested that participants discussed the study with their parents and, therefore, instructed participants to take their briefing and consent forms home for counter-signing by their parents. Counter-signing their daughter's form was taken to indicate parental consent for participation. One participant failed to return a counter-signed consent form (despite a number of reminders) and she was asked politely to withdraw from the study.
In keeping with guidance from Frankfort-Nachmias and Nachmias (1992, cited in Cohen et al., 2008), I also took steps to maintain participant confidentiality. Each participant was instructed to follow a procedure that assigned them a unique, anonymous pseudonym – the first two letters of their mother's maiden name and the last two digits of their phone number. This strategy preserved the identity of participants, but allowed completed instruments to be paired with their respective respondents during different phases of data collection, I chose to keep this data for 2yrs (as recommended by Sieber, 1998, cited in Creswell, 2003).

However, whilst the ethical issues mentioned so far are relatively germane across educational research, two additional ethical questions required special consideration. Namely: a) is it justifiable to expose participants knowingly to a significant academic stressor and, b) how to recruit participants in a school where, not only am I known to the students, but hold a position of authority within the school's senior management team.

In regard to the second point, in their Code of Conduct, the British Psychological Society emphasise that;

"Investigators should realise that they are often in a position of authority or influence over participants... [and that] this relationship must not be allowed to pressurise the participants to take part in, or remain in an investigation" (BPS website).

Therefore, although my role within the school gave me arguably greater authority as a researcher in this study than that of other researchers in
other projects, it is a problem common to all research scenarios, rather than an ethical point specific to this study. Therefore, I took steps similar to that of other researchers to recruit participants; namely the use of a neutral incentive (see recruitment section below).

The other ethical concern - ‘is it justifiable to expose participants knowingly to a stressor?’ - was addressed in two ways. Firstly, by deliberately utilising a pre-existing academic stressor (modular public examinations in the January 2013 season), which would have taken place irrespective of the students’ participation in this study. As such, it is likely that the study generated little additional stress and, therefore, that the risk/benefit value of the study was extremely low (Sieber & Stanley, 1988). Secondly, in keeping with literature documenting the positive impact of debriefing in reducing stress\(^{56}\) (e.g. Robinson & Mitchell, 1993; 1995), all participants underwent a debriefing in groups of 5 - 8 at the end of the study “to remove any harmful effects brought on by the study, and to leave participants with a sense of dignity and a perception that their time was not wasted” (Harris, 1988, pg191). In keeping with guidance from Aguinis and Henle (2004), the debriefing included a summary of the research findings, an opportunity express opinions and thoughts about the study and a chance to ask questions.

Recruitment. In order to recruit participants I advertised the study in three ways: 1) through announcements in two school assemblies, 2) by giving brief introductory talks on the biochemistry of stress during Sixth

\(^{56}\) There is some debate as to the efficacy of debriefing in reducing chronic, or post-traumatic stress. However, recent reviews (e.g. Everly & Mitchell, 2000) uphold its role in stress reduction.
Form Biology and Chemistry lessons and, 3) by giving presentations outlining the research study in Sixth Form tutor group meetings. This approach was rationalised by research from Patel, Doku and Tennakoon (2003, pg234), who state that “few participants will take part in research unless they can identify with and understand its validity and relevance.” Additionally, as many of the participants were completing A-level courses in one or more science subjects, or who aspired to study a science-related degree at university, I also incentivised the study as an opportunity for students to participate in a ‘real research study’. This is in keeping with advice offered by Hoinville and Jowell (1978, cited in Cohen et al., 2008, pg110) that incentives should be “clearly seen as a token, rather than a payment for the respondent’s efforts and should be as neutral as possible.”

This incentive proved particularly appealing (a fact emphasised by the dozen or so girls who mentioned ‘participating in a real research study’ in their UCAS personal statements) and, in total, 119 female Sixth Formers volunteered to participate in the study. However, as four students failed to complete all three phases of data collection and a further student withdrew for ethical reasons (her parents had not counter-signed her consent form), only 114 participants contributed data towards the completed study. This formed a Control group of 44 participants and an Experimental group of 70 students.

**Non-random grouping.** Despite the non-randomised nature of the quasi-experimental protocol, grouping participants in this manner came as close to arbitrary allocation as was possible under the circumstances. Briefly, seven curriculum departments (Mathematics, Geography, Biology,
Chemistry, Politics, Economics and Philosophy & Ethics) have it as policy that students studying for these courses sit examinations in the January examination season, both for the AS and A2 year groups. These subjects are relatively diverse and include a wide spectrum of students' aptitudes, ambitions and interests; thus lending significantly towards heterogeneity of grouping. Furthermore, as students sitting the January examinations had not elected to do so, but instead had been entered as an indirect consequence of selecting their A-level subject options, the confounding influence of academic ability (or GI) upon group selection was reduced. This was further assisted by the precaution of excluding participants from the Experimental group if they were only re-taking AS examinations - i.e. only students sitting AS or A2 modules for the first time were selected for the Experimental group.

**Statistical Validity.** Likert-type instruments are known to produce interval-type data because "[although] the response categories have a rank order, the intervals between values cannot be presumed to be equal" (Jamieson, 2004, pg1217 - see also Appendix 2). However, despite this, large scale meta-analyses observe that parametric statistical tests are widely employed in methodologies incorporating Likert-type instruments (Clason & Dormody, 1994\(^{57}\)), resulting in Kurzon to decry the use of parametric statistics on ordinal data, arguing that this oversight constitutes the first of "seven deadly sins of statistical analysis" (Kurzon et al., 1996, pg265). However, whilst Jameson's point is well-made, others argue that the use of parametric testing on Likert-generated data is justified. This argument is supported in two distinct ways; 1) the

\(^{57}\) Clason and Dormody (1994) analysed 95 studies and found that only 13% used nonparametric measures.
contention that, whilst single Likert items should be considered ordinal, Likert instruments consisting of sums across many items (such as those employed here) are effectively interval (Carifio & Perla, 2008). 2) the observation that, even if the 'wrong' statistical test is used, often the outcome of either test leads towards the same conclusion: “this is what statisticians call robustness - the extent to which the test will give the right answer even when assumptions are violated” (Norman, 2010, pg626). The robustness of parametric tests as applied to Likert-type instruments is well evidenced: for example, meta-analyses frequently demonstrate that parametric and nonparametric statistical tests yield similar power (Clason & Dormody, 1994; De Winter & Dodou, 2012), whilst other studies have found that, in circumstances where responses are relatively normally distributed, similar (low) incidences of Type I between test regimes (De Winter & Dodou, 2012 - and never more than 3% above the nominal rate of 5% [Lipsey & Hurley, 2009]). Thus, as Norman (2010) asserts:

"Parametric statistics can be used with Likert data, with small sample sizes, with unequal variances, and with non-normal distributions, with no fear of 'coming to the wrong conclusion'. These findings are consistent with empirical literature dating back nearly 80 years" (Norman, 2010, pg632).

The design and validity of the instruments

As stated earlier, this study utilises three specific instruments - the PSS10, the ESES and the SA-IAT. Of these (bar some very minor adjustments - discussed below) two of these instruments are, effectively, 'off the peg' (the PSS10 and the ESES) and, therefore, bring with them a wide body of prior validatory research. By contrast, though
heavily based on a cogent prototype, the absence of a pre-existing instrument necessitated the construction the SA-IAT myself. This process (and the obligatory trialling and validation steps associated with the design phase) are conferred in detail in my Institution Focused Study (see Appendix 2) and are, therefore, not discussed any further here. However, before detailing individually the evidence that supports the use of these specific instruments in this project, because they are commonly united by their use of self-report methodology, I wish briefly to outline my position on the validity of self-report measures per se.

Traditionally, procedures for establishing the validity of the hypothetical, psychological constructs self-reports aim to measure tended to centre around establishing unequivocal, empirical observations that demonstrated Construct Validity - i.e. the degree to which "a test measures what it claims, or purports, to be measuring" (Brown, 1996, pg231). Methodologies for establishing construct validity tended to assess whether "one's measure of a given construct related to measures of other constructs in a theoretically predictable way" (Smith, 2005, pg395), effectively justifying or disproving the validity of one construct based on empirical evidence gathered from the measurement of another. Though popular in the 1950s and '60s, such polarized 'justificationism' (Bartley, 1962) has more recently fallen out of favour, with authors such as Lakatos (1999) concluding that there is now a significantly greater appreciation for the 'indeterminate and ongoing nature of theory building and scientific criticism.'

In relation to self-reports: "the psychological processes underlying an act of self-reporting are understood to be exceedingly complex" (Paulhaus &
Vazire, 2006, pg224) and, because multiple factors may contribute to the content of individual items (Kessler, 1987), some authors assert (e.g. Paulhaus, 1991) that it is rarely clear precisely what construct self-reports actually measure.

"At the most basic level, there is concern about the construct validity of self-report measures. Both theory and research indicate that self-report responses are a product of psychological, sociological, linguistic, experiential and contextual variables, which may have little to do with the construct of interest" (Razavi, 2012 pg4).

Thus, in contrast to the original assertions of Cronbach, as theories and techniques for knowledge assessment have evolved, so too has our understanding of construct validity (Smith, 2005) in that research no longer strives to demonstrate unequivocal validity, but rather to view the validation process as an on-going dialogue between evolving theories and methodologies. In the following section, I draw together elements of this dialogue, which support both my decision to employ self-report measures in this project, and also outline my reasoning for choosing specifically the PSS10, SA-IAT and ESES instruments.

**Examination Self-Efficacy Scale.** Traditionally, SE has been measured quantitatively using a self-report instrument and, over the last three decades, a large number of such SE-assessment instruments have been developed (e.g. Scherer's [1982] Self-Efficacy Scale and the New General Self-Efficacy Scale developed by Chen and his colleagues [Chen et al., 2001]). Of these instruments, however, the General Self-Efficacy Scale (GSES) has emerged most popular, chiefly because it is free to use and
open-source, as well as being thoroughly validated in a wide range of research projects (e.g. Scholz et al., 2002). It is, therefore, the obvious instrument to use as a starting point in this research.

Developed in Germany and translated into English by Schwarzer in 1995 (Schwarzer & Jerusalem, 1995) the GSES was originally constructed as an assessment of General Self-Efficacy. However, this construct has been heavily criticised (e.g. Stajkovic & Luthans, 1998) as it runs in stark contrast to Bandura's original concept of domain-specific SE (Bandura, 1997). Thus, as “SE is typically defined as [a person's] perceived capabilities within specific domains” (Schunk & Pajares, 2002, pg27 - my emphasis), the GSES, whilst a logical starting point for the measurement of SE, had to be adapted in order to measure specifically academic SE\(^5\). This precaution was intended to avoid measurement attenuation through generality with other SE domains. In fact, one of the secondary appeals of using the GSES is that, as Schwarzer asserts, the GSES “can be easily adapted to tap specific behaviour [by] add[ing] or change[ing] a few items to cover the particular content of the survey” (Schwarzer, 2011). Therefore, to incorporate the changes that altered the GSES to become a specific measure of academic SE, I refer to the adapted version of the GSES as the Examination Self-Efficacy Scale (ESES – see Appendix D) from this point onwards.

Duly adapted, the ESES now consists of 20 items, each of which scores between 1 and 6 points utilising a Likert-type scale. Questions 1, 3, 4, 7,

\(^5\) Note that even across separate branches of academic SE the generality of SE appears limited (see Bong, 1999; Pajares, 1996 & Ferrari & Parker, 1992).
9, 12, 15, 16, 18 & 20 are ESES questions and are scored from 1 - 6, whereas the other ten questions are fillers and attract no score. Thus ESES scores range from a minimum of 10 to a maximum of 60. The inclusion of ten 'filler' questions was designed to obfuscate response direction - i.e. to limit the degree to which participants were able to manipulate their responses towards a specific outcome (a reported criticism of questionnaire methodology - see Williams, 2003) and also to avoid what Haladyna (1997, pg11) describes as Proximity Error, which can occur “when differing skills are rated similarly when sequentially ordered [in a questionnaire].” Additionally, to overcome Acquiescence Bias (the tendency to say 'Yes' indiscriminately in answer to questions [Toner, 1987]) sequential items were adapted to elicit both positive and negative responses i.e. to balance the GSES; for example, I changed Question 2 to include the word 'not' ("I do not know what I want to do in the future) so that respondents are not tempted to bias their answers through indiscriminate positive responses. The use of a balanced Likert-type scale also controls Leniency and Severity Errors (the tendency to give high ratings to most items by agreeing / disagreeing with everything) as, with half of the questions adapted to elicit opposite answers to the other half, errors of these types are effectively averaged out across the items.

Likert and Likert-type questionnaires are designed to comprise unidimensional items, which “are of equal value in that they each provide a replicated assessment of the dimension measured by the total score on the scale” (Siniscalco & Auriat, 2005, pg61). Assessments of internal variance of the GSES consistently report Cronbach's alpha of between .70 and .90 (Schwarzer, Mueller & Greenglass, 1999; Dona et al., 2002), with “most [alpha] scores greater than .80” (Schawrzer, 2011).
Therefore, by basing the ESES on the format of the GSES, the instrument measures "a continuous underlying dimension assessed by total scores on the attitude scale" (Siniscalco & Auriat, 2005, pg60).

As well as utilising unidimensional items, Likert-type scales have also traditionally employed a five- or six-item response scale as Johns (2010, pg6) asserts that "data from Likert items... becomes significantly less accurate when the number of scale points drops below five or above seven." Whilst Johnson, Smith & Tucker (1982) confirm that skew increases for Likert-type scales with three scale-point responses or less, Dawes (2002) found little difference between response accuracy in 5- and 11-point scales, indicating that Likert-type items with higher numbers of response points might produce equally accurate data. This observation was similarly demonstrated for kurtosis by Dawes (2008), who found comparably (low) levels of this kind of central tendency bias between 5-, 7- and 11-point responses to Likert-type items. That said, in the interest of 'manageability' Johns (2010, pg6) advocates utilising a smaller response scale where possible, arguing that "few people will have a clear idea of the difference between, say, the eighth and ninth point on an 11-point agree-disagree scale." Therefore, as Schwarzer has generally allocated 10min for completion of the GSES (Schwarzer & Hallum, 2008), adapting GSES to utilise a larger response scale would not only take respondents longer (and, potentially, generate disincentive and a lower completion rate), but would be unlikely to result in any net gain in accuracy, nor loss of skew nor reduction in kurtosis. Therefore, I adapted the ESES to utilise the smaller six-category response scale.
By tradition, Likert scales include an odd number of response items so that respondents are offered a neutral midpoint (the 'neither agree nor disagree' response), which helps identify indifferent and ambivalent responses from respondents whose opinions are genuinely directed, but not strong enough to access the stronger response categories at the ends of the scale (Nowlis et al., 2002). However, in keeping with Schwarzer's (2011) original design, the ESES has only six semantic response categories, thus forcing participants to bias their response in one direction or another. Johns (2005) advocates such 'Likert-type' scales (i.e. not a true Likert scale as it lacks a central neutral response 'fulcrum' or 'seesaw' [McCall, 2001]) because "some people use a midpoint to avoid reporting what they see as less socially unacceptable answers" (Johns, 2005, pg238). His research into odd- and even- Likert-type scales demonstrated that respondents answer more positively to questionnaires that use an odd number of scale points - "the midpoint attracts many of those who actually disagreed, but were reluctant to admit as much" (Johns, 2005, pg239). The same evidence is replicated by Garland (1991, pg68), who concluded that "resorting to a scale without a mid-point seems to alleviate social desirability bias without changing the direction of opinion." Therefore, although Worcester and Burns (1975) point out that smaller response scales can alter the intensity of respondents' opinions, the limiting effect of a six-category response on the Social Desirability Bias makes it the preferred option for this questionnaire. Additionally, Shaw and Wright (1967) posited that responses in the middle category were frequent when respondents either had no clear opinion, or had not yet formed a clear opinion; thus:
“The middle response category... is at least occasionally utilized as a 'dumping ground' for not applicable, uncertain, indifferent or ambivalent response orientations” (Kulas & Stachowski, 2009, pg492).

Therefore, the ESES has been designed to lack a neutral response position and, because of this, it constitutes a more accurate gauge of SE than odd-numbered response scales.

Schwarzer (2011) does not provide a justification for his choice of ten scored items for the GSES. However, Gilem & Gilem (2003) advise against smaller questionnaires, favouring multi-item measures as the most appropriate tool for assessing psychological attributes. They assert that questionnaires with low numbers of items (or even a single item) have considerable measurement error (Gilem & Gilem, 2003), which must be averaged across a larger number of responses to increase reliability. They also advocate a multi-item scale “to discriminate among fine degrees of an attribute” (Gilem & Gilem, 2003, pg8). By contrast Herzog and Bachman contend that;

“Survey instruments have a maximum length beyond which there is an increasing probability of premature termination, random responding, or other behaviour patterns which result in data of lower quality” (Herzog and Bachman, 1981, pg558)

Their conclusion was that shorter questionnaires (which they loosely define as 45min or less) show less Position Bias than questionnaires with larger numbers of items. Johnson and his colleagues agree (Johnson et al.,
1974), and further advocate shorter questionnaires on the grounds of limited response omission, as their study clearly demonstrated that questions placed within the first 18 items of a questionnaire had significantly higher response levels compared to questionnaires in where the same question appeared later on. Therefore, the choice of twenty items for the ESES is in keeping with this recommendation, whilst ensuring that the instrument is long enough both to overcome measurement error and to assess academic SE discriminatively, but not so long as to generate bias or response omissions.

Apropos of omitted responses, Schwarzer recommends “our rule of thumb is to calculate a score as long as no more than three items on the ten-item scale are missing” (Schwarzer, 2011, pg3). Though something of a moot point (as no participant returned an incomplete instrument), had instruments been returned incomplete, I planned in this instance not to follow Schwarzer’s guidance and instead to remove the entire dataset from subsequent statistical analysis. This precaution would have lessened the erosive impact of omitted responses on internal validity, whilst also reducing compound measurement error.

In addition to providing evidence of reliability, “establishing Stability is also vital to validating psychometric tools… Stability refers to the concept that constructs retain a degree of resistance to change over time” (Lane et al., 2004, pg 339). Across a number of test-retest studies SE self-reports have been recorded as having high stability coefficients, with a gap of weeks (Luszczynska et al., 2005; Leganger et al., 2000), months (Schwarzer et al, 1999) and even up to two years (Chen et al., 2001).
One final source of error germane to questionnaires is *Halo Effect Error*, which Siniscalco and Auriat (2005, pg6) describe as "the tendency to rate a particular statement according to how respondents feel about it in general." Cooper (1981) identifies four potential sources of halo: engulfing, insufficient concreteness, under-sampling & insufficient participant motivation. The latter sources are overcome respectively by including an appropriately large number of questionnaire items and allowing respondents enough time to complete the questionnaire. I have already addressed these issues previously. According to Cooper (1981, pg222) engulfing occurs when responses "are coloured by an overall impression" such that "judgements on seemingly unrelated dimensions are engulfed by a general impression" (Feeley, 2002, pg226). However, as the ESES constitutes a unidimensional assessment of a specific form of academic SE, engulfing is, in fact, the desired outcome of the questionnaire! In other words, I wish responses to individual items to generate (or 'be engulfed by') an overall impression of academic SE. By contrast, however, Cooper's fourth source of halo - *Insufficient Concreteness* - is relevant to Likert-type questionnaires and occurs when items lack specificity or appear vague or ambiguous. Feeley's (2002, pg234-235) advice here is to "use concrete, specific and clear categories to assess instruction" and to "define items clearly and precisely and include an example to clarify the item to respondents." Therefore, to reduce or overcome halo, I included exemplars in the participant briefings and also altered syntax in the ESES to be as deliberate and clear as possible.
However, despite the strengths of the ESES questionnaire and its practical utility, there are some limitations of Likert-type methodology. A major restriction is the use of discrete response categories which, by nature of their ‘coarse granularity’ (Chimi & Russell, 2009), can at best only be viewed as an ordinal qualitative scale i.e. it cannot be assumed that one person’s ‘moderately true’ is the same as another’s and “there is no assumption of equal intervals between categories” (Cohen et al., 2008, pg327). However, as the SA-IAT instrument has exactly the same limitation (see Appendix B), this is a moot point for the ESES. A second limitation is that there is no way of assessing whether participants have responded truthfully, even with the preventative bias-reducing measures outlined above, there is no absolute certainty that the data collected denotes a ‘true’ representation of the respondents’ academic SE. A third limitation is that “often the total score of a respondent has little clear meaning since a given total score can be secured by a variety of answer patterns” (Kothari, 2008, pg83). Therefore, despite the unidimensional nature of the ESES, I cannot assume that participants with similar ESES scores have identical SE beliefs. I return to this discussion in Chapter 6.

**Perceived Stress Scale.** The 10-item Perceived Stress Scale (PSS10 – see Cohen & Williamson, 1988) is a shorter and more widely used version of the initial PSS self-report instrument conceived by Cohen, Kamarck and Mermelstein (1983). Both versions of the PSS are used “to evaluate the degree to which situations in one’s life are appraised as stressful” (Cohen et al., 1983, pg385). Thus the PSS constitutes “a subjective estimate of global stress level” (Cohen & Williamson, 1988, pg34) and, as such, the PSS has become “among the most commonly used self-report..."
measure of perceived stress in the social sciences” (Lavoie & Douglas, 2012, pg48).

The PSS10 (see Appendix E) is a Likert instrument similar in design to the ESES. Participants respond using a 5-point scale ranging from 0 (never) to 4 (very often). Of the 10 items, 4 items (4, 5, 7 & 8) are worded in a positive direction and were, therefore, reverse-scored. As with the ESES, the inclusion of such reversed items helped to balance the instrument, which in known to limit acquiescence bias (Toner, 1987) and to reduce leniency and severity errors. Similarly, the choice of employing the shorter PSS10 instrument in preference to longer alternatives is in keeping with the rationale for the design of the ESES in that it should help to limit the development of position bias across the instrument (Johnson et al., 1974). One stark difference between the design of the ESES and the PSS10 is that the PSS10 makes use of an uneven response scale, which is at odds with my earlier assertion that scales with a fulcrum, or midpoint, tend to encourage social desirability bias (Johns, 2005) and, therefore, are prone to kurtosis. On balance, although I would have preferred to alter the instrument to include a sixth response category (as was the case with the ESES), the existence of a significant body of validitatory studies (see below) and the opportunity to utilise normative values (Cohen & Janicki-Deverts, 2012), support a decision to employ the PSS10 as designed (Cohen & Williamson, 1988). Thus the PSS10 was not altered and retained its original five-point response scale.

As indicated, the validity of the PSS10 has been tested extensively. For example, assessments of internal validity consistently report high alpha scores (e.g. .89 - Roberti et al., 2006; .85 - Cohen et al., 1993; .78 Cohen
analyses have been unable to identify any meaningful differential item-bias by sex, race, education (Cole, 1999), age, employment status or income (Cohen & Janicki-Deverts, 2012). Similarly, factor analyses reliably demonstrate that the PSS10 constitutes a bidimensional measure of stress (Cohen & Williamson, 1988; Gitchel et al., 2011; Ramírez & Hernández, 2007; Roberti et al., 2006) in that it assesses ‘perceived distress’ and ‘perceived coping ability’ as distinct factors which, when summated, constitute an internally valid perceived global measure of stress (Cohen & Williamson, 1988; Hewitt et al., 1992; Lavoie & Douglas, 2012). This conclusion is further supported by confirmatory evidence from test-retest studies (Cohen et al., 1983) and the consistency and volume of reports demonstrating the reliability of the PSS10 in a wide variety of different countries and across a diverse range of cultural contexts. For example, the PSS10 has been translated into at least twenty different languages (Cohen, 2012) including; Spanish (Remor & Carrobles, 2001), Swedish (Eskin & Parr, 1996), Chinese (Lee & Crocket, 1994) and Japanese (Mimura & Griffiths, 2008), as well as the original English (Cohen & Williamson, 1988). Each study reports effects that support Cohen’s assertion that the PSS10 constitutes “a universal measure of perceived stress” (Cohen et al., 1983, pg386).

Similar to reliability, the stability of the PSS10 has also been extensively tested, with authors publishing a high degree of concordance across tests separated by a year (Golden-Kreutz et al., 2004), or over repeated measures throughout that period (Byrne et al., 1989).
In addition to its consistent stability and demonstrable content validity, the PSS10 also exhibits a high level of construct validity. For example, convergence has been demonstrated with the State-Trait Anxiety Inventory (Robert et al., 2006), the Stress Vulnerability Scale (Connor et al., 2007), the Stress in Academic Life Scale (Alzaeem et al., 2010), the Daily Stress Inventory (Machulda et al., 1998) and the Multidimensional Health Locus of Control (Robert et al., 2006) as well as a number of other measures of stress (Cohen & Williamson, 1988) and depression (Mimura & Griffiths, 2008). Similarly, divergence has been observed with instruments that do not measure stress (Cohen & Williamson, 1988), including the Sensation Seeking Scale, the Santa Clara Strength of Religious Faith Questionnaire and the Overt Agression Scale (Robert et al., 2006). Further evidence of construct validity comes from studies in which participants with high self-reported stress levels (using the PSS10 instrument) demonstrated a correlating elevated level of salivary cortisol (Baker et al., 1984; Van Eck & Nicholson, 1994; Malarkey et al., 1995), thereby showing concordance with physiological measures of stress.\footnote{It should be noted that this association has not been observed in all studies (e.g. Ebrecht et al., 2004; Van Eck et al., 1996) possibly because of differences in testing methodologies (Stowell, 2003; Bosch et al., 2004), such as ambiguity in the measurement of acute vs chronic stress (Murphy et al., 2010).}, which further strengthens the construct validity of the PSS10.

Despite Cohen’s (2012, pg2) warning that “the Perceived Stress Scale is not a diagnostic instrument” and his insistence that the predictive validity of the PSS “is expected to fall off rapidly after four to eight weeks” (Cohen et al., 1997, pg11) because “levels of appraised stress should be influenced by daily hassles, major events, and changes in coping resources” (Cohen, 2012, pg2), a growing number of studies have
established strong predictive validity of the PSS10 with a diverse range of pathologies, including infertility (Band et al., 1998), skin conditions (Chiu et al., 2003), suppressed immune functioning (Kiecolt-Glaser et al., 1995), depression (Otto et al., 1997; Treadgold, 1999) and ageing (Epel et al., 2004), to mention but a few (see also Cohen & Janicki-Deverts, 2012; Cohen, 2012). Additionally, a number of studies highlight the PSS10 as predictive of maladaptive stress-related behaviours, such as smoking (Cohen et al., 1993), missing sleep and consuming alcohol (Cohen & Williamson, 1988), which may begin to explain the strength of the connection between high PSS10 scores and stress-related illness. Either way, whether the link between PSS10 scores and stress-related morbidity (Nielsen et al., 2008) is directly causal, or whether it represents an indirect correlation through maladaptive behaviours (Kiecolt-Glaser & Glaser, 1988), the conclusion must be that the PSS10 constitutes a valid measure of stress in that it is strongly predictive of stress-related illness.

Therefore, as the PSS10 demonstrates robustness in all three subcategories of validity as identified by Messick (1989) - i.e. content validity, construct validity and predictive validity - I conclude that the PSS10 constitutes a valid measure of global stress (Cohen & Williamson, 1988). That said, there are some limitations with the PSS10, of which the most significant is:

"The potential influence of other variables such as the subject's personality, mood during testing and psychopathology. Furthermore, if these types of variables are associated with a psychological disorder (diagnosed or undiagnosed), then the subject's ability to
appraise their stress is suspect, thereby nullifying the instrument’s predictive validity” (Herbert & Cohen, 1996, pg300).

As with the limitations of the ESES, I return to this discussion in Chapter 6.

**Stressor-Appraisal Implicit Association Test.** Implicit Association Tests (IATs) provide a quantitative assessment of the relative strength of automatic appraisals using a reaction-time paradigm (Greenwald et al., 1998). The IAT measures these automatic, or implicit evaluations by:

“Assessing the strength of mental associations between a target concept (e.g. ‘exams’) and one pole of an evaluative dimension (e.g. ‘stressful’), as compared to a contrast concept (e.g. ‘holidays’) and the opposite pole of the evaluative dimension (e.g. ‘relaxing’)” (Andrews et al., 2010, pg2388).

Participants are asked to categorize stimulus words into a series of different target-concept groups as quickly as possible. The speed of their reaction time is measured and provides an indirect representation of the preconscious association between the individual target and the individual concept. Overall, therefore, IATs work on the assumption that “it ought to be easier to make the same behavioural response [a key press] to concepts that are strongly associated, than to concepts that are weakly associated” (Nosek et al., 2005, pg167). Thus we might expect the key-press timings of two strongly associated concepts (e.g. ‘exams’ and ‘stressful’) to be faster than the key-press timings of two weakly associated concepts (e.g. ‘holidays and stressful’). By presenting the initial
target categories (e.g. ‘exams’ and ‘holidays’) and the associated concept attributes (e.g. ‘stressful’ and ‘relaxing’) in all possible combinations, the underlying relative implicit associations between these concepts can be evaluated.

In this research, I employ an IAT specifically designed to measure the preconscious association between stress and exams - the Stressor Appraisal Implicit Association Test (SA-IAT). This instrument produces scores based on Greenwald's D measure (Greenwald et al., 2003), which is an effect-size-like measure with a possible range of zero to two around a ‘neutral’ position of one. Scores less than one denote a stronger association between ‘holidays’ and ‘stressful’, whereas scores higher than one represent associations between ‘stressful’ and ‘exams’. As the design and validation of this instrument formed the focus of my Institution Focused Study (which I enclose in full as the second appendix to this thesis), instead of reiterating that research here, I simply restate my conclusion from that research that the SA-IAT constitutes a valid measure of implicit stressor appraisal. I also urge the reader to read Appendix 2 for a full evaluation of the SA-IAT and a justification of my assertion that the SA-IAT instrument is an appropriate tool for use in this research project.

Précis of the methodology

Thus far I have explained the rationale for adopting a quasi-experimental methodology and presented justifications for implementing an untreated control group design with one pre-test. I have proffered summaries of the individual instrument protocols and contextualised their use in this
study with statements of validity. I also have outlined the control procedures and processes employed to increase the validity of the methodology, and given an overview of the ethical considerations of this study, and the steps taken to ensure conformity with guidelines. With the methodological narration now complete, in the next Chapter I present the research findings from the study and details of the statistical tests utilised that I employed in order to reach those findings.
Chapter 5: Research Findings

This Chapter summarises the statistical measures taken to address the three research questions, the outcomes of those processes and limitations to the significance of conclusions. I also take steps to confirm statistically the validity of the instruments.

A summative table of data analysed in this Chapter is presented in Table 1 and 2 below.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>PrS Testing</th>
<th>PoS Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SA-IAT</td>
<td>ESES</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Means and SDs for PrS & PoS data for the Control group.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>PrS Testing</th>
<th>PoS Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SA-IAT</td>
<td>ESES</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Means and SDs for PrS & PoS data for the Experimental group.

Addressing the research questions

1. To what extent does academic SE limit the development of acute academic stress? A simple two-tailed student's t-test for independent samples was conducted to compare Control (M=14.10, σ=2.50) and Experimental (M=17.11, σ=2.22) PoS PSS10 data. An initial Levene's Test
revealed equivalent variances ($F[1,112] = .676, p = .413$) and, therefore, a $t$ statistic assuming homogeneity of variance was computed, which revealed a significant difference in reported stress between the two participant groups ($t[112] = -6.375, p < .001$). This can be seen clearly in Figure 4, where error bars (here representing 95% confidence limits) clearly exhibit no overlap. This suggests that there was a significant difference in self-reported acute academic stress levels between Control and Experimental participant groups subsequent to academic stressor exposure.

![Figure 4: Average PrS & PoS PSS10 scores for Experimental and Control groups.](image)

However, whilst the $t$-test confirms a difference in PoS data, it fails to consider both the non-randomised nature of the grouping procedure and the PrS level of stress (which, in fact, was also significantly different
between groups – see footnote\(^{60}\). Therefore, in order to compensate for these considerations, a one-way ANCOVA analysis was conducted incorporating PrS PSS10 scores as a covariate and the type 3 Sums of Squares algorithm to compensate for uneven group sizes (George & Mallery, 2003). The ANCOVA (between-subjects factor: PoS PSS10; covariate: PrS PSS10) also revealed a significant difference in PoS reported academic stress levels between groups (F[1,110]=126.78, \(p < 0.01, \eta^2_p=.533\)), which confirms the conclusion suggested by earlier t-test data – i.e. that participants in the Experimental group experienced significantly more acute academic stress than those in Control group. Additionally, the ANCOVA also highlighted the effect of stress experienced prior to academic stressor exposure, as PrS PSS10 scores were also found to have a significant effect on PoS PSS10 responses (F[1,110]=111.09, \(p < 0.01, \eta^2_p=.500\)). Finally, the interaction between pre-stress and group variables was found to be insignificant (F[1,110]=.122, \(p=.728, \eta^2_p < .001\), thus demonstrating that the regression slopes for PrS PSS10 between groups are homogenous (see Figure 5 below), which ensures that the ANCOVA procedure was appropriate for the data collected.

\(^{60}\) A student's t-test for independent samples (t[112]=2.281, \(p=.34\)) compared PrS data across participant groups. It revealed a significant difference between PSS10 group scores (Control: \(M=15.70, \sigma=2.48\); Experimental: \(M=14.76, \sigma=1.93\)) before stressor exposure.
Having established a significant difference in PoS PSS10 levels between groups, the relationship between academic SE and PoS acute academic stress was examined. To that end, an ANCOVA was conducted (between-subjects factor: PoS ESES; covariate: PrS ESES) to determine whether the participant groups had significantly different academic SE. The results of the analysis confirmed no significant difference between group academic SE ($F[1,110]=.021, p=.886, \eta^2_p < .001$) and no significant interaction between group and PrS academic SE ($F[1,110] < 0.001, p > 0.99, \eta^2_p < .001$ – see Figure 6). Thus, pre-stress academic SE was the significant factor in determining post-stress academic SE ($F[1,110]=518.49, p < 0.01, \eta^2_p=.825$).

Assuming equivalence of academic SE, the correlation between PoS ESES and ΔPSS10 data (the difference between PrS and PoS PSS10 scores)
was calculated for each participant group using Pearson's Product Moment Correlation Coefficient (Control: $r[42]=-.127$, $p=.413$; Experimental: $r[68]=-.585$, $p < .001$ - correlations are represented in Figure 7). Correlations were analysed using a Fisher's r-to-z transformation calculation (Lowry, 2001), which revealed a significant difference ($Z=2.74$, $p < 0.01$) between correlations. However, as the ESES-$\Delta$PSS10 correlation used in the transformation calculation was not significant, the outcome of this test must be interpreted with caution.

![Figure 6: Homogeneity of regression for ESES data for Control and Experimental groups.](image-url)
To address the limitations of the previous test and to gauge quantitatively the contribution of academic SE in the genesis of acute academic stress, two linear multiple regression analyses were conducted (one for each participant group). The regressions assessed the extent to which PoS PSS10 data (dependent variable) could be predicted from PrS PSS10 data and ESES scores (unrelated independent variables). For the Control group, the regression equation was significant ($R^2=.596$, adjusted $R^2=.576$, $F[2,41]=30.254$, $p<0.01$) and revealed a significant contribution from pre-existing stress (PrS PSS10: $t[41]=5.875$, $p<0.01$) and a very nearly significant contribution from academic SE (ESES: $t[41]=-1.999$, $p=0.52$). Standardised coefficients revealed that 64.9% of stress was predicted by pre-existing stress and -22.1% was attributable to academic SE. This effect was considerably pronounced in the Experimental group, where -47.7% of academic stress was predicted by academic SE, whilst the contribution of pre-stressor exposure stress was found to be consistent.

Figure 7: Correlation between ESES and ΔPSS10 for Control and Experimental groups.
with the Control group at 62.6%. The regression equation was also significant for the Experimental group ($R^2=.684$, adjusted $R^2=.674$, $F[2,67]=72.369$, $p < 0.01$), although here the contribution from both variables was significant (PrS PSS10: $t[67]=9.058$, $p < 0.01$; ESES: $t[67]=-6.895$, $p < 0.01$)

Therefore, in summary, evidence from the statistical tests described here reveals that exposure to an academic stressor stimulus resulted in an increased level of self-reported academic stress in the Experimental group. This effect was limited substantially by the participants' perceived SE. Furthermore, evidence from the Fisher's r-to-z transformation suggests that the stress-reducing effect of SE was more pronounced for participants with greater pre-existing SE than those with average, or below-average levels of academic SE (although this effect was not established within established confidence limits).

2. To what extent does acute academic stress exposure affect the preconscious appraisal of academic stressor stimuli? Two two-tailed student's $t$-tests for independent samples were conducted to compare similarity of SA-IAT between Control and Experimental groups in both the PrS and PoS phases of testing. In both cases Levene's Tests revealed equivalent variances (PrS: $F[1,112]=.607$, $p=.438$; PoS $F[1,122]=.502$, $p=.472$), so $t$ statistics assuming homogeneity of variance were computed. For the first test PrS data was used (Control: $M=1.092$, $\sigma=.143$; Experimental: $M=1.104$, $\sigma=.163$) and no significant difference between Control and Experimental group SA-IAT scores was identified ($t[112]=-0.400$, $p=.690$). This can be clearly seen in Figure 8, where error bars for the Control group (displaying 95% confidence limits) demonstrate
considerable overlap. By contrast, error bars for the Experimental and Control PoS groups share virtually no overlap at all, implying a significant difference between scores. This was confirmed by the second $t$ test, which found a marked and significant difference between these data sets ($t[112]=-4.418, p<.001$).


due to the effect of PrS artefacts and non-randomised grouping on PoS SA-IAT scores, a one-way ANCOVA analysis was also conducted. The ANCOVA included PrS SA-IAT scores as a covariate and utilised the type 3 Sums of Squares algorithm to compensate for the uneven group sizes (George & Mallery, 2003). The ANCOVA (between-subjects factor: PoS SA-IAT; covariate: PrS SA-IAT) also revealed a significant difference in Control and Experimental PoS SA-IAT scores ($F[1,111]=30.795, p<0.01, \eta_p^2=.217$), confirming that the Experimental group did indeed demonstrate a stronger association.
between the target category 'exams' and the concept attribute 'stressful'. As with the previous ANCOVA, homogeneity of regression was established by observing no significant interaction between Group and PrS SA-IAT scores ($F[1,110]=.958, p=.330, \eta^2=.009$ - see also Figure 9).

![Figure 9: Homogeneity of regression for SA-IAT data for Control and Experimental groups.](image)

The histograms shown in Figure 10 compare the distribution of pre- and post-stressor exposure SA-IAT scores. The elevation in PoS SA-IAT scores is clearly visible.
3. To what extent does academic SE affect the preconscious appraisal of academic stressor stimuli? Four Pearson's Product Moment Correlation Coefficient tests were calculated to assess the correlation between academic SE and SA-IAT score for the Control and experimental participant groups in both the PrS and PoS phases. The correlations were: PrS - Control: \( r[42]=-0.082, p=0.598 \); Experimental: \( r[68]=-0.095, p=0.434 \); PoS - Control: \( r[42]=0.194, p=0.208 \); Experimental: \( r[68]=-0.201, p=0.095 \). The correlations are shown in Figure 11.

Figure 10: Histograms of SAI-AT Scores showing the difference in distribution between Control (top) and Experimental (bottom) group in PrS (left) and PoS (right) phases.
Correlations were analysed using two Fisher’s r-to-z transformation calculations (Lowry, 2001), which revealed no significant difference (Z=.07, p=.472) between PrS correlations and a significant difference (Z=.2.02, p=.022) between PoS correlations. However, as all four initial correlations used in the transformation calculations were not significant, the outcome of this test must be interpreted with caution.

To assess whether the Fisher’s r-to-z transformations had revealed a significant correlation between academic SE and SA-IAT scores, a two-tailed student’s t-tests for independent samples was conducted to compare the similarity of ESES scores between two subpopulations of the Experimental group - those who had experienced a drop in SA-IAT scores between phases of testing (i.e. a decrease in the extent to which stressors were appraised as stressful; N=21, M=-0.057, σ=.038) and those who had experienced an increase (N=49, M=0.128, σ=.104). An initial
Levene’s Test revealed equivalent variances \((F[1,69]=.291, p=.591)\), so \(t\) statistics assuming homogeneity of variance were computed. The student’s \(t\)-test revealed a significant difference in ESES scores between sub-populations \((t[68]=2.272, p=.026)\), demonstrating that participants in the Experimental group with higher academic SE had indeed appraised stressors less stressful than participants with lower academic SE. This relationships is also visible in Figure 12 below.

![Figure 12: Difference in SE scores between subpopulations of the Experimental group. Left bar represents population with negative ΔSA-IAT scores (PoS-PrS), right bar represents population with positive ΔSA-IAT scores.](image)

To gage quantitatively the contribution of academic SE in the implicit appraisal of stressor stimuli, linear multiple regression analyses were conducted for each participant group to assess the extent to which PoS SA-IAT data (dependent variable) could be predicted from PrS SA-IAT data and ESES scores (unrelated independent variables). For the Control group, the regression equation was significant \((R^2=.606, \text{ adjusted } R^2=.587,\)
F[2,41]=31.526, \( p < 0.01 \) and revealed a significant contribution from pre-stress SA-IAT score (PrS SA-IAT: \( t[41]=7.645, \ p < 0.01 \)) and an insignificant contribution from academic SE (ESES: \( t[41]=1.238, \ p=0.223 \)), which had a standardised coefficient of 12.2%. By contrast, for the Experimental group, the contribution of academic SE in predicting SA-IAT scores was found to be significant (ESES: \( t[67]=-2.635, \ p=.01 \)), with SE contributing to -21.9% of variation. The pre-stress SA-IAT score was found to predict 72.3% of SA-IAT score (PrS SA-IAT: \( t[67]=8.758, \ p < .01 \)). The regression equation was also significant for the Experimental group (\( R^2=.542, \ \text{adjusted} \ R^2=.529, \ F[2,67]=39.707, \ p < 0.01 \)).

Therefore, in summary, whilst academic SE was not found to play a significant contribution to the implicit appraisal of academic stressor stimuli for the Control group. By contrast, for the Experimental group, participants with low academic SE exhibited significantly elevated SA-IATs after exposure to the academic stressor, whilst participants with higher academic SE did not.

Validity of the instruments

The internal validity of IAT instruments is well documented (Banse et al., 2001; Bosson et al., 2000; for a review see Nosek et al., 2007), both through studies demonstrating high alpha scores (e.g. Steffens & Buckner, 2003) and by research establishing robust temporal stability (Egloff et al., 2005). Whilst the omission of a second phase of pre-testing precluded the assessment of the internal validity of the SA-IAT in this study, I was able to conduct a small post-study test-retest procedure following a method similar to that used by Dasgupta and
Greenwald (2001). Thirty-five participants from the Control group repeated the SA-IAT 24hrs after PoS testing. The correlation between first and second SA-IAT scores was examined using Spearman's Rho and this measure of reliability was then corroborated by calculating Cronbach's alpha. As Figure 13 shows, test-retest data were significantly correlated ($r_{[35]}=.705, p < .001$), indicating a level of reliability similar to that reported by Bosson and colleagues (Bosson et al., 2000) and a level of reliability higher than that of many other studies (e.g. Greenwald & Farnham, 2000; Dasgupta & Greenwald, 2000). Alpha calculations ($\alpha=.833$) also support a conclusion that the SA-IAT is a reliable instrument.

![Figure 13: Test-retest reliability for the SA-IAT.](image)

The reliability of the other instruments (the ESES and the PSS10) was also assessed by calculating Cronbach's alpha; although, in contrast to the SA-IAT, calculations for these instruments were based on data collected during PrS testing. In keeping with the conclusions of previous authors
(e.g. Schwarzer et al., 1999; Dona et al., 2002; Siniscalco & Auriat, 2005),
the ESES was found to be highly reliable (α=.896). By contrast, the
PSS10 demonstrated a lower internal consistency (α=.661) than that
reported by most authors (e.g. Roberti et al., 2006; Cohen et al., 1993;
Cohen & Janicki-Deverts, 2012), although a number of studies have
published comparable alpha values (Ramadoss & Bose, 2010 -.70;
Wongpakaran & Wongpakaran, 2010 -.71; Cohen & Williamson, 1988 -.78).
Therefore, despite the relatively low PSS10 alpha score, in this research
the PSS10 displayed an 'acceptable' level of reliability, whilst the SA-IAT
and ESES respectively demonstrated 'good' to near 'outstanding'
reliability (George & Mallery, 2003).

In addition to demonstrating their internal validity, all three instruments
elicited response profiles very similar to published norms, thus
demonstrating convergence with established standards and, therefore, a
high degree of construct validity. For example, all three instruments
achieved levels of significance greater than .05 in Shapiro-Wilk tests for
both Control and Experimental groups (SA-IAT: $\text{Sig}(\text{Cont})=0.381$,
$\text{Sig}(\text{Exp})=0.172$; PSS10: $\text{Sig}(\text{Cont})=0.67$, $\text{Sig}(\text{Exp})=0.125$; ESES: $\text{Sig}(\text{Cont})=0.293$,
$\text{Sig}(\text{Exp})=0.068$). Similarly, Normal Q-Q plots (an example of which is given
in Figure 14) demonstrated a high level of concordance with regression
lines which, taken in conjunction with the results from normality tests,
established that the data from all three instruments was normally
distributed - the accepted profile for these tests. Furthermore, mean
PrS PSS10 scores for both Control and Experimental groups (Control:
$M=15.7$, $N=44$, $\sigma=2.48$; Experimental $M=14.8$, $N=70$, $\sigma=1.93$) were well
within half a standard deviation of norms published for females under 25
years old (Cohen & Janicki-Deverts, 2012). Similarly, ESES means were
found to be very similar to normative values (Schwarzer, 2011) and, in keeping with data from previous studies (e.g. Greenwald et al., 2009; Sriram & Greenwald, 2009), SA-IAT responses were identified as following a normal distribution with PrS means very close to the 1.00 neutral position (Control: M=1.09; Experimental: M=1.04 – see also Figure 9). This was the case for both participant groups.

![Normal Q-Q Plots for PrS SA-IAT data for Experimental group.](image)

**Figure 14**: Normal Q-Q Plots for PrS SA-IAT data for Experimental group.

Distribution analyses confirmed that data from ESES responses and SA-IATs displayed no significant kurtosis for either participant group (both PrS and PoS) and, across all response items, data for the ESES was found to be free from skew. Although slight negative skew was identified in the PrS SA-IAT data ($S_K=-0.783$), the $S_K$ value was considerably less than the critical value of ±1.96 (Sheskin, 2011) and was, therefore, not significant. By contrast, the low alpha score in the PSS10 was

---

61 Note that this incorporates an adjustment of Schwarzer’s (2011) normative values to accommodate adapting the instrument to include a 6-interval response scale.
accompanied by mild, though significant, negative skew in half of the PSS10 response items (questions 1, 2, 6, 9 & 10 elicited $S_k$ values > 1.96). Additionally, significant leptokurtic distribution was also identified in three PSS10 response items (questions 2, 8 & 10 elicited kurtosis scores > 1.96). I discuss these limitations in more detail in Chapter 6.

**Review of the findings**

In this Chapter I have described the statistical procedures employed to analyse the data collected throughout the study. I have stated, where appropriate, the confidence limits of these measures and applied this to test outcomes to highlight where significance was achieved. In sum, the outcomes of these processes are encapsulated as follows;

**To what extent does acute academic SE limit the development of academic stress?**

1. Prior to academic stressor exposure, there was no significant difference in academic stress levels between participant groups.

2. After exposure to the academic stressor, the Experimental group reported more acute academic stress than the Control group.

3. There was no difference in academic SE between participants in the Control group and the Experimental group.

4. The Experimental group showed a significant negative correlation between academic SE and the degree of acute academic stress experienced.
5. Academic SE predicted nearly 50% of this effect.

To what extent does acute academic stress exposure affect the preconscious appraisal of academic stressor stimuli?

1. Prior to academic stressor exposure, there was no significant difference in implicit academic stressor appraisal scores between participant groups.

2. After exposure to the academic stressor, the Experimental group reported elevated implicit academic stressor appraisal scores.

To what extent does academic SE affect the preconscious appraisal of academic stressor stimuli?

1. Academic SE predicted approximately -20% of the increase in implicit stressor appraisal scores.

In the next Chapter I discuss the ramifications of these findings, draw conclusions from the observations described in this Chapter, discuss the limitations of this research and speculate on future research directions.
Chapter 6: Discussion

Individuality of stress

In keeping with prior research (e.g. Lohaus, 1990; Lakshmi, 2009; Chamberlain et al., 2011), this study confirms the findings of others that exposure to academic stressor stimuli constitutes a potent elicitor of acute academic stress. In this study the Experimental group (M=17.11, σ=2.22) reported significantly greater levels of acute academic stress than the Control group (M=14.10, σ=2.50) in post-stressor exposure testing (see Figure 15), but not in pre-stressor exposure testing, implying a causal relation between academic stressor exposure and the genesis of acute academic stress.

Figure 15: Post-stress PSS10 scores for Control group (above) & Experimental group (below).
This association is further confirmed by the frequency of PSS10 scores of 20 or above, which a number of authors have used to differentiate between acute stress and acute distress (e.g. Stauder et al., 2009; Wilks, 2009). In this study only one member of the Control group reported such distress (2.3%) and their score remained the same in both phases of testing (thus, I speculate, the distress was likely to have had a social underpinning). By contrast, however, 11 members of the Experimental group (15.7%) reported post-stressor exposure PSS10 scores of 20 or more, where only two had reported distress in the previous phase of testing. Furthermore, in both of these cases, the level of distress increased further after academic stressor exposure, supporting the same conclusion - i.e. that exposure to academic stressors contributes significantly to the acute academic stress experienced by adolescents. Given the evidence I presented in Chapter 2 demonstrating that chronic stress either predisposes (or potentially causes) adolescents to develop psychopathologies (e.g. Seiffge-Krenke, 2000), disease (e.g. Ohlin et al., 2004; Rosengren et al., 2004) and increased morbidity (Nielsen et al., 2008), the implications of my conclusion are concerning, particularly as I posit that some of the participants in this study are at some risk of developing chronic stress as a consequence of further academic stressor exposure after the end of the study (e.g. through additional public examinations).

However, despite observing significant acute academic stress in many of the participants, it is both interesting and reassuring to note that 40% of students in the Experimental group reported academic stress levels lower than the normative mean for their age and gender (Cohen & Janicki-Deverets, 2012 - see reference line on Figure 15), despite being exposed
to the same academic stressor stimuli that had made many of the other participants so distressed. This outcome is absolutely in keeping with the observations of others (e.g. Williams et al., 2011), who repeatedly report similar wide variations in stress levels when different people are exposed to the same stressor. Not only does this reliable finding offer some degree of hope in averting chronic stress becoming “our next public health crisis” (Anderson, 2010, pg17), but it also serves to reiterate the question that ultimately inspired me to begin this study - why do people react so differently to similar stressors?

Effectively, this observation both highlights and reinforces the well-established view that stress is highly individualistic (Cooper & Bright, 2001) and greatly subjective (Lazarus & Erikson, 1952; Lazarus, 1999). Indeed, the large standard deviation of post-stressor exposure PSS10 scores reported in this study (and that of others - Cohen & Williamson, 1988; Cohen et al., 1993; Cohen & Janicki-Deverts, 2012) clearly supports Lazarus and Folkman’s (1984, pg19) assertion that stress is, fundamentally, “based on personal perception and individuality.”

To what extent does academic SE limit the development of acute academic stress?

In his Cognitive Mediational theory, Lazarus (1966; 1968) explains this individuality of stress in terms of the cognitive interpretation of stressors. Such evaluations - or appraisals (Lazarus, 2001; 1999; 1993) - are one of 9 sources of evidence I outlined in Chapter 2 to support viewing stress as an affective process. Among the ‘cognitive resources’ (Sherman et al., 2008) identified as contributing towards the conscious
appraisal of stressors, SE has been found to play a significant role. Specifically, SE has been shown to protect people from the effects of stress (Bandura, 1997) as numerous studies have demonstrated that SE is negatively correlated with stress (e.g. Betoret, 2006; Vaezi & Fallah, 2011). In this study, the first research question sought to test the degree of generality of this finding, by investigating the correlation between academic SE and academic stress in a group of female adolescents. Although SE is known to develop with age (Anderman et al., 1999; Urdan & Midgley, 2003), the link between SE and reduced stress is well-established by late adolescence (Hackett et al., 1992; Solberg et al., 1993; Gigliotti & Huff, 1995; Solberg & Villarreal, 1997; Torres & Solberg, 2001), which enabled me to adapt instruments and methodologies used in prior research with adults (e.g. Schwarzer, 2011) for use in this study.

In this project the Control group displayed no visible correlation between academic SE and self-reported academic stress - as expected; whereas the Experimental group was found to show a significant and strongly negative correlation between academic SE and acute academic stress after exposure to an academic stressor stimulus (see Figure 7). This effect was found to be independent both of the degree of pre-existing stress and also of the variation in academic SE between groups. Thus the answer to the first research question is that academic SE was found to limit the development of acute academic stress in the participants involved in this study. Furthermore, although I was unable to establish that the difference in SE-stress correlation was significant between participant groups, regression analyses suggested that academic SE may ‘protect’ participants by removing as much as half of the effect of
academic stressor exposure in the Experimental group. This effect is in keeping with that found by other studies (e.g. Zajacova et al., 2005 - where SE was predictive of stress to a similar extent).

However, by themselves, such correlations do not constitute evidence of causality between academic SE and the development of acute academic stress; nor do they suggest a mechanism by which this protection might occur - criticisms which both apply equally to my study. However, Bandura (1997) has collated evidence from numerous studies supporting a causal role for SE in protecting people from stress, the majority of which I have summarised earlier in Chapter 3. Of this evidence, the most evincing comes from school-based stress-reduction programmes, which have clearly established that enhancing young people’s academic SE results in decreased stress (e.g. Kraag et al., 2006), both in class-based schemes (Lohaus, 2011; Hampel et al., 2008; Frydenberg, 2004) and outdoor-based activities (Thurber et al., 2007; McKenzie, 2000; Masten & Coatsworth, 1998). Not only do these prospective, experimental studies employ methodologies that most strongly support causality (Shadish et al., 2002), but they are also most ecologically valid (Messick, 1995) as they provide in situ, real-time evidence of academic SE limiting the acute stress experienced by young people on a day-to-day basis. Thus, not only do these studies strongly support a causal link between academic SE and acute stress reduction (which, I posit, generalises particularly to this study through cogent context), but I suggest they also offer the most promising route forwards for policy-makers and educational leaders in helping to equip young people with an element of protection against future exposure to academic stressors.
Whilst the weight of evidence supports the conclusion that academic SE does indeed protect adolescents from acute academic stress, identifying the mechanism by which this happens is much less straightforward. As I explained in Chapter 3, most research to date attempts to explain the link between SE and stress in terms of adaptive changes in coping style (e.g. Taylor, 1989; Rothbaum et al., 1982; Brandtstädtler, 1992; Lazarus & Folkman, 1984; Lazarus, 1999). Although evidence exists to support a coping-based explanation (e.g. Mikolajczak et al., 2007), maladaptive changes in coping style, by themselves, are not always predictive of the development of chronic stress (Chung et al., 2001), or the occurrence of stress-related diseases (Blalock & Joiner, 2000), suggesting that explanations of the protective nature of SE couched in terms of adaptive coping alone might not hold all of the answers. Thus I proposed in Chapter 3 that SE may also exert an effect through the appraisal of stressors and that this effect may be more significant for the genesis of stress, rather than the regulation of pre-existing stress. Whilst evidence exists to support a role for SE in mediating the conscious appraisal of stressors (Chemers et al., 2001; Lazarus & Folkman, 1984; Pintrich & De Groot, 1990), I have been unable to identify any research seeking to establish whether SE exerts a similar effect on preconscious appraisal processes. Therefore, the third research question sought to fill this 'gap' in our understanding of stress by investigating the effect of academic SE on the preconscious appraisal of academic stressor stimuli. However, in order to address this question, I first had to establish the role of preconscious appraisal processes during acute academic stress. This, therefore, formed the basis of the second research question.
To what extent does acute academic stress exposure affect the preconscious appraisal of academic stressor stimuli?

During acute stress, such as that experienced by the participants in the Experimental group in this study, it is well established that preconscious appraisals bias conscious attention towards threatening stimuli (e.g. Rodrigues et al., 2009; LeDoux, 1998), generating the well-observed ‘attentional tunnel’ phenomenon (see Staal, 2004). This adaptive response (De Kloet et al., 1999) occurs jointly through repression of neocortical functioning (specifically the PFC - Elzinga & Roelofs, 2005; Oei et al., 2006) and elevations in the sensitivity of preconscious appraisal systems to stressors (specifically the amygdala - Van Marle et al., 2009). The combined effect of this activity acts to facilitate the encoding of stressor-specific memory (Zoladz et al., 2011) and, crucially, to divert conscious attentional resources towards the stressor stimulus (Bar-Haim et al., 2007).

In this study I sought to test the degree of generality of this finding by observing whether preconscious stressor appraisals changed during the course of the acute academic stress process experienced by the participants in the Experimental group. As expected, whilst I found no significant differences in preconscious stressor appraisals before exposure to academic stressor stimuli (in either of the participant groups); by contrast, after exposure, the Experimental group demonstrated a significant increase in SA-IAT scores (see Figures 8 &

62 Although note that this occurs only for stress-specific ‘flasbulb’ memories as acute stress has routinely been shown to have a general inhibitory effect on the formation of other memories (e.g. Diamond et al., 1996)
10), whereas the Control group’s scores remained relatively unchanged. Thus the answer to the second research question is that acute academic stress exposure appeared to affect the preconscious appraisal of academic stressor stimuli in the participants involved in this study. In other words, the experience of acute academic stress resulted in some participants evaluating academic stressors as being more stressful than they would otherwise have appeared. This observation was found to be independent both of the initial variation in SA-IAT scores and also of pre-existing academic stress. Furthermore, it can be readily explained using Chee’s theory of IAT effects (Chee et al., 2000): according to this model, the inhibitory effect of acute stress on cognitive processes during conditions of acute stress necessitates greater conscious effort to override automatic responses to stressors, which produces a greater delay in overall response – i.e. the IAT effect observed here.

Therefore, the second finding of this study fits perfectly with the pre-existing literature by demonstrating that exposure to acute academic stress resulted in a significant increase in the degree to which adolescents preconsciously appraised academic stressors as being stressful. However, whilst this conclusion holds true for the average SA-IAT scores between participant groups, an analysis of the scores of specific participants reveals a very different pattern for many of the participants. For example, Figure 16 clearly shows that a marked proportion (30%) of the Experimental group exhibited either little change or a decrease in their SA-IAT scores after experiencing acute academic stress. Thus, although the Experimental group demonstrated, on average, a significant increase in the degree to which stressor stimuli were preconsciously appraised as stressful, this effect was not
homogeneous. This is not surprising given the observation that, despite exposure to significant daily stressors (e.g. road traffic accidents), some people neither experience symptoms of acute stress (Paton & Violanti, 2006) nor go on to develop stress-related diseases (Prati et al, 2010).

Clearly the line of inquiry to take forward from this observation is why did some one third of the participants in the Experimental group fail to show the same trend as the majority? In other words, what prevented them from developing the same degree of academic stress as their peers?

Having demonstrated previously that SE exerts a significant ‘protective’ effect in limiting the development of stress, I posed the third research question to ascertain whether academic SE might explain this trend – i.e.

Figure 16: Distribution of variation in SA-IAT scores between test phases (PoS-PrS) for Control group (above) & Experimental group (below).
whether the 'protective' role of SE might be mediated via changes in the preconscious appraisal of stressors.

**To what extent does academic SE affect the preconscious appraisal of academic stressor stimuli?**

Comparing the difference in average academic SE between sub-populations of the Experimental group (those who showed *increases* in SA-IAT score between testing phases and those who showed *decreases*) revealed a significant difference between groups. This suggests that participants with lower academic SE tended to appraise academic stressors as being *more stressful* after acute academic stress exposure. This effect was further supported by the analysis of correlations in academic SE and SA-IAT scores between participant groups in both pre- and post-stressor exposure phases of testing. Furthermore, regression analyses found that academic SE was predictive of -22% of variance in SA-IAT scores, implying that high academic SE provides much of its protective effect by limiting, or *buffering*, the up-regulating effect of acute academic stress on the preconscious appraisal of academic stressor stimuli.

Thus, the answer to the third research question is that academic SE appeared to limit the degree to which academic stressor stimuli were preconsciously appraised as being stressful. Although considerable additional work is required to establish whether this effect is causal (and not merely, for example, a correlation contingent on an indirect relationship between these variables); nevertheless, this study has provided the first evidence to suggest that academic SE exerts at least
part of its protective, stress-limiting role via an effect on preconscious appraisal processes.

Although this conclusion is firmly in keeping with literature demonstrating the protective role of SE in ameliorating stress (e.g. Skaalvik & Skaalvik, 2007; Goddard et al., 2000; Hoy & Spero, 2005), it could appear to run at odds with Bandura’s (1997) conception of efficacious human agency with intentionality and forethought at its heart (Bandura, 1997; 2008). However, although Bandura views agentic function as ontologically emergent - i.e. an epiphenomenal product of human consciousness (Bandura, 1991b; 2006b), considerable evidence exists to suggest that this agentic “measure of control over... thoughts, feelings and actions” (Pajares, 2002, pg2) is as much a product of implicit processes as it is of conscious thought. For example, research using behavioural priming methodology has repeatedly shown that social behaviour can be instigated automatically by environmental stimuli (Dijksterhuis & Van Knippenberg, 1998), or that people will unintentionally adopt primed physical behaviours - the ‘chameleon effect’ - even in circumstances where there is no incentive to do so (Chartrand & Bargh, 1999). However, the observation that goal-directed behaviours can also be activated automatically (e.g. Chartrand & Bargh, 2002; see also Bargh et al., 2010) and result in the same judgmental outcomes as consciously derived behaviours (Bargh et al., 2001), constitutes clear evidence that purposive agentic behaviour may just as easily stem from preconscious processes as it does from explicit cognition. Indeed, recent findings from the field of automaticity now suggest that preconscious processes may be

---

63 These are two of Bandura’s four ‘core beliefs of agency,’ the others being self-reactiveness and self-reflectiveness (Bandura, 1997; 2008).
capable of judgement and decision making (Dijksterhuis & Nordgren, 2006; see also Bargh et al., 2012) - processes long considered the "last bastions of conscious processing" (Bargh, 2011, pg632-633). Thus, “in both phylogeny and ontogeny, actions of an unconscious mind precede the arrival of a conscious mind” (Bargh & Morsella, 2008, pg73 - abstract) and, with ready access to explicit beliefs, values and goals, preconscious systems enable us to “influence the course of events by one’s action” (Bandura, 2011, pg52) and thus function as efficacious human agents. Therefore, in the words of John Bargh:

“Conscious thought is causal and it often puts automatic processes into play: similarly, automatic processes regularly cause and influence conscious thought processes. These two fundamental forms of human information processing work together, hand in glove, and indeed one would not be able to function without the support and guidance of the other” (Bargh et al., 2012, pg602).

Thus I assert that human agency interacts with both conscious and preconscious appraisal processes and that, in the light of this study, both systems of stressor appraisal appear to be influenced by our efficacy beliefs.

Next Steps

The findings of this study, whilst enthrusting, leave a large number of unresolved issues outstanding that urgently require investigation. However, rather than listing lots of these here, I draw attention specifically to three further research questions, which I believe form the next priorities for research in this area.
1. **By what mechanism/s does SE interact with preconscious appraisal processes?** It is currently unclear whether SE exerts the effect identified in this study by interacting directly with preconscious systems, or via an indirect effect through modulation of neocortical appraisal processes and resultant alteration in 'top-down' influence on automatic processes. As I stated in Chapter 3, this study appears to be the first project to research potential interactivity between SE and preconscious appraisal processes and, therefore, further conjecture into the nature of the relationship would simply be pure speculation. Clearly, further work is required to unpick the nature of the observations recorded here, perhaps through the use of induced magnetic lesion methodology to remove top-down inhibition from neocortical systems? Similarly, advances in temporal neuroimaging techniques may also help to identify the time-course of brain activity that underpins the link between SE and preconscious stressor appraisals? Either way, the findings of this study supports mounting evidence that:

"Emotional information processing at an early level of processing, with limited conscious awareness, represents an important step in regulating the magnitude of... stress" (Ellenbogen et al., 2010, pg79).

I look forward to further developments in this field which, I hope, will shed light on this important and potentially life-saving field of stress research.

2. **By what mechanism/s does chronic stress interact with preconscious appraisal processes?** Deliberately, this study has focused
specifically on adolescents’ acute response to academic stressors and the interaction both academic SE and the stress process itself appear to exert on it. However, as the majority of the deleterious physiological and psychological repercussions of stress detailed in Chapter 2 occur as a consequence of chronic stressor exposure, rather than initial acute stress; it is, therefore, imperative that research is also conducted in the field of chronic stress in order to establish whether preconscious stressor appraisals are equivalently malleable in this process.

In Chapter 2 I asserted that stress and affect share functional equivalence; given this position, it follows that research into the regulation of emotion may inform our understanding of the potential adaptability of preconscious appraisals that, I speculate, may occur during chronic stress. Within AER, evidence is beginning to accumulate to suggest that pathological failures of emotion regulation (such as avoidant personality styles) can be explained through disruption of top-down regulatory systems. For example, Phillips and colleagues (Phillips et al., 2008) have suggested that neocortical dysregulation engenders either disinhibition of preconscious valence detection, or increased preconscious monitoring of affective state (Philips et al., 2008), or possibly both. Collectively, the resulting effect is to unbalance preconscious appraisals, which generates a maladaptive attentional bias towards emotive stimuli and, therefore, a predisposition toward psychoses. Could it be that a similar dysregulated ‘top down’ effect underpins the chronic pathology of stress?\textsuperscript{64}

\textsuperscript{64} It should be noted that, outside of stress research, theories highlighting the significance of maladaptive preconscious appraisal processes in the development of psychopathologies are plentiful. For example, dual processes models of addiction (Gladwin et al., 2011; Wiers et al., 2007; Bechara, 2005) identify that automatic systems
Within the field of chronic stress research, it is well known that preconscious appraisal processes may start to regulate cognitive attention adaptively by diverting attention away from the stressor stimuli (Ellenbogen et al., 2009; 2010). For example, Fabes and colleagues found that participants reporting high stress levels were more likely to disengage attention from other stressor stimuli (Fabes et al., 1993). Similarly, Mansell adapted a dot-probe task to incorporate social stressor words and observed that participants with high trait anxiety were significantly more likely to avoid threat cues than participants with low anxiety (Mansell et al., 1999) – a finding that was later replicated by Applehans and Lueck (2006). Comparable effects of stress-induced attentional bias have also been reported in modified spatial cueing tasks (Ellenbogen et al., 2002; 2009), Stroop-type tasks (Van Honk et al., 2000) and during the Trier Social Stress Test (Pilgrim et al., 2010).

Based on this view, Lueck (2006) have advanced a model for maladaptive chronic stress that exhibits congruence with those proposed for AER. Their theory is based on the findings of a number of studies (e.g. Mogg et al., 2000; Wilson & MacLeod, 2003; Huguet et al., 1999) that, under some circumstances, during conditions of prolonged stressor exposure, preconscious appraisals can become maladaptive. For example, Chajut and Alg (2003) used a Stroop-type paradigm to show

underpin many of the drug-related behaviours of addicts (Ryan, 2002; Wiers & Stacy, 2006). Similarly, research into anxiety disorders (Bar-Haim et al., 2007; Mogg & Bradley, 1998; Cisler & Koster, 2011; Öhman, 2000), depression (Gotlib et al., 2004), bipolar disorder (Philips et al., 2008), Post-Traumatic Stress Disorder (Harvey et al., 1996), Panic Disorder (Schmidt & Woolaway-Bickel, 2006) and a variety of Eating Disorders (Faunce, 2002) also emphasises the key role biases in preconscious appraisal systems play in the development of many mental illnesses.
that manipulating task-related stress resulted in *increased* attention towards stressor stimuli during chronic stress. This effect was replicated by Roelofs and colleagues (Roelofs et al., 2007), who observed that preconscious regulation of conscious attention was bidirectional — i.e. the allocation of attention was contingent on magnitude of the cortisol response to prior stressors. Luecken and Appelhans (2005) suggest that this effect can be explained by stress-induced disruption of top-down neocortical systems during chronic stress. Normally, during chronic stress, these areas suppress the preconscious appraisal bias towards the stressor, resulting in the attentional avoidance reported in studies like Mansell’s (Mansell et al., 1999). However, in pathological conditions, this process becomes disrupted, resulting in amplification of the preconscious appraisal bias in a manner very similar to that reported during anxiety, depression and addiction. This theory is well supported, both by the reliable observation that exogenous cortisol injections dampen the threat-related preconscious attentional bias (Putman et al., 2007; 2010) and the wealth of evidence demonstrating the harmful effects of cortisol on neocortical regulatory functioning (McEwen, 2007). However, to date, this theory remains exactly that — a theory — as experimental data has proven conflicting. For example, Ellenbogen and colleagues (Ellenbogen et al., 2002; 2006; 2010) have employed stimulus masking paradigms to show that disengagement from stressors can occur independently of the behaviour of conscious appraisal processes. Similarly, the activity of these neocortical systems is not always predictive of HPA reactivity during chronic stress (Van Honk et al., 2000; Ellenbogen et al., 2006; 2010), which suggests that chronic stress exerts its effect on preconscious appraisals via direct modulation of these processes. Thus it is currently unclear whether chronic stress exhibits its effect on
preconscious systems through a direct effect, via indirect top-down modulation or preconscious appraisals, or possibly by means of both mechanisms (Mogg et al., 2004; Pine et al., 2005). As I stated earlier, further work is required in this area to address this area of uncertainty, particularly given the “exorbitant toll” (Matheny et al., 1993, pg109) chronic stress exacts each year from young people across the UK.

3. To what extent does academic SE buffer the preconscious appraisal of stressors during chronic stress? A very large number of studies exist which clearly demonstrate the stress-limiting effect of SE in adolescents. Collectively, these studies assess stress across a myriad of different contexts (e.g. Holden, 1991; Holden et al., 1990), including interpersonal stress (Rumi & Kunio, 2003); academic stress (Solberg et al., 1993; Solberg & Villarreal, 1997; Torres & Solberg, 2001) and social stress (Dalgard et al., 2012), to name but a few. However, despite the compelling unanimity of their conclusions, as with this study, the majority of such investigations have tended to focus on the role of SE in the acute stress response. By comparison, however, comparatively little research has investigated the effects of SE on chronic stress. Studies which have investigated chronic stress tend to focus either on the links between SE and stress-related disease (e.g. O'Leary & Brown, 1995; Caserta et al., 2011) or stress-induced psychopathologies (e.g. McFarlane et al., 1995), most particularly anxiety disorders (Matsuo & Arai, 1998; Yue, 1996) and depression (Cutrona & Troutman, 1986). Although some research has investigated the effect SE appears to play in mediating the chronic stress response (e.g. Buddeberg-Fischer et al., 2000; Beehr & Newman, 1978), such investigations almost exclusively involve adult participants and frequently centre around prolonged exposure to workplace stressors.
Therefore, not only is there a pressing need to investigate the effect of academic SE on chronic stress experienced by adolescents (i.e. within schools), but, given the conclusions of this study, it is also crucial to investigate whether, during conditions of chronic stress, SE exerts a similar protective to that documented in this study by buffering preconscious stressor appraisals.

Conclusions

In summary, the conclusions of this study are as follows:

1. Stress is an inherently subjective process.

2. Academic SE exerts a significant protective effect in limiting the development of acute academic stress.

3. For adolescents with lower academic SE, exposure to acute academic stress increases the degree to which academic stressor stimuli are preconsciously appraised as stressful.

4. For adolescents with higher academic SE, exposure to acute academic stress appears not affect the degree to which academic stressor stimuli are preconsciously appraised as stressful.

Implications of the conclusions

In the opening Chapter of this study I made the point that efficacy theories offer significant potential to develop pedagogy, policy and
practice that could successfully reduce the “exorbitant toll exacted from students by stressors” (Matheny et al., 1993, pg109). Having demonstrated in this study that academic SE effectively protects adolescents from the effects of an acute bout of academic stress (via a buffering effect on preconscious stressor appraisal processes), I again assert that efficacy beliefs offer a valuable and effective vehicle to combat the rising incidence and increasingly harmful effects of stress in young people (Twenge et al., 2010).

Limitations of the conclusions

Limitations of the methodology. Whilst I have justified my decision to employ a quasi-experimental design in this study (see Chapter 4), the use of this methodology comes with inherent limitations. Of these, “the lack of random assignment is the major weakness of the quasi-experimental study design” (Harris et al., 2006, pg17), chiefly because the lack of randomisation has an implicit tendency to introduce confounding variables within the participant groups. For example, although general intelligence (GI) is considered not to affect the development of stress (Singh & Sharma, 2012), it was not possible to control entirely for this potentially confounding factor across the participant groups. Although I took steps to limit the effect of GI on grouping demograph (by preventing students who were re-taking AS units only from participating); nevertheless, the influence of re-taking AS examinations will inevitably have skewed the level of GI across the two participant groups. Similarly, although I argued earlier that the seven subjects offering January units were “relatively diverse and include a wide spectrum of students’ aptitudes, ambitions and interests,” they are nonetheless biased towards the Sciences and Social
Sciences and, most notably, do not include representation from either the creative arts, or modern languages. As it has been shown that levels of stress can differ substantially across different academic disciplines (Clark, 1997) it seems plausible that the differences in academic pursuit that shaped the participant groups, may also have obfuscated accurate measurement of the dependent variable.

In a similar vein, it is possible that the use of non-randomised participant groupings introduced non-academic confounding variables into the participant groupings. For example, fatigue is well-known to predispose adolescents to stress (e.g. Akerstedt et al., 2002) and, because it is also associated with diminished academic performance (Wolfson & Carskadon, 2003), attenuated concentration, loss of motivation, increased anxiety and a host of other direct and indirect effects on emotion regulation and cognition (Dahl, 1999), discrepancies between the degree of rest experienced by members of the different participant groups over the 2012 Christmas vacation may well have affected the distribution of stress described in this study. To explain, because members of the Control group did not have any forthcoming public examinations in the January 2013 examination season, it is likely that they undertook significantly less independent preparatory study over the vacation than did members of the Experimental group. As a consequence of this reduced workload, it is likely that the Control group enjoyed significantly more rest and recuperation over the Christmas vacation than did members of the Experimental group who, by nature of their revision workload, were likely to have experienced less sleep, or to have spent less time enjoying leisure activities. Thus, imbalances in convalescence between participant groups may well have resulted in differences in
fatigue levels amongst the participants, which could potentially have biased the outcomes of PrS testing and, *ita sequitur*, significantly reduced the external validity of my conclusions.

In addition to differences in fatigue, when compared with members of the Experimental group, participants in the Control group are also more likely to have spent time away on holiday over the Christmas vacation. As this factor is well-associated with reductions in stress (Pols & Kroon, 2007), the opportunity (or absence of opportunity) to go away on holiday constitutes another uncontrolled factor associated with the use of non-randomised participant groups in this quasi-experimental methodology that may, potentially, have affected the findings of this study.

Collectively, Shadish describes the effect of such uncontrolled factors as History Threats (Shadish et al., 2002), asserting that such threats “greatly limit the external validity of [quasi-experimental] procedures” (Cook & Campbell, 1979, pg45). Although the history threats I have described so far occurred as a consequence of the use of non-randomised participant groups, the existence of uncontrolled factors may well have biased experimental outcomes independently of the grouping protocol. For example, *personality* is predictive of health (Friedman, 1990) and psychopathology (Contrada et al., 1990) and a wide body of research demonstrates that specific personality variables, or *traits*, either seem to provide people with resilience against stress (e.g. De Neve & Cooper, 1998; Carver et al., 1993), or may predispose them to it. For example, participants scoring highly on measures of trait anxiety are reliably shown to be susceptible to developing stress (Spielberger, 1985). Similarly, negative affect (Watson & Clark, 1984), neuroticism
(Schneider, 2004), Type A personality (Rosenman et al., 1964), Type D personality (Denollet, 2000) and a variety of other personality types are well-correlated with stress and the development of stress-related disease. In this study I took no steps to assess either the level, or the effect of these personality traits within the participants, so it is conceivable that the differences in PrS PSS10 scores I attributed to academic SE, in fact owed to variation within the distribution of personality traits across the participants. As well as personality type, specific dispositional characteristics of personality have also been implicated in the development of stress (e.g. aspiration [Kadapatti & Vijayalaxmi, 2012], self-expectation [Angola & Ongori, 2009] and expectation from family members [Ang & Huan, 2006] also promote stress), all of which are potential confounding variables, which could conceivably have led to the pattern of stress described in this study.

As well as personality, lifestyle and environmental factors are also heavily implicated in the development of stress: smoking (Parrott, 1999), lack of exercise (Ensel & Lin, 2004), diet (e.g. Raloff, 2000) and increased coffee consumption (Pincomb et al., 1987) have all been shown to promote stress in some circumstances. By contrast, time spent socialising (House et al., 1988), moderate alcohol consumption (Sayette, 1999; although see Pohorecky, 1991) and the presence of a family pet (Siegel, 2010) have been implicated in stress reduction. Similar to personality traits, these factors were neither controlled, nor statistically regressed in this study, and may well have biased the results reported earlier. Directly linked with lifestyle, socioeconomic status is strongly predictive of chronic stress (e.g. Baum et al., 1999) and the development of stress-related illness (Adler et al., 1994), chiefly because it predisposes adolescents
towards lifestyle choices that are associated with the development of stress (Cohen et al., 2006). Although I argued earlier that, being an “independent and academically selective fee-paying day school, the influence of many of the other social factors known to affect SE (e.g. socioeconomic status…) was limited,” this factor was, nevertheless, uncontrolled in this study, and may well reduce the external validity of the conclusions presented earlier.

In keeping with the limitations presented earlier, the largest category of history threat in this study pertains to coping theory. Simply put, coping can be defined as:

“Conscious volitional efforts to regulate emotion, cognition, behaviour, physiology, and the environment in response to stressful events or circumstances” (Compas et al., 2001, pg89).

Despite the vast body of literature that exists in this field, and the diverse domains of coping that stem from it (see Skinner et al., 2003; Snyder, 1999; Lazarus, 1993b), most researchers agree that:

“The set of adaptive processes that give… [people] the potential to fend off disaster, to reshape challenges and to transform stressful experiences into psychological growth” (Skinner, 2007, pg245).

Thus coping makes a huge difference both on the impact stressors exert on adolescents and also on the likelihood of ensuing stress resulting in illnesses and chronic medical conditions, both concurrently and long-term (Skinner, 2007). Although this study focused on the implicit appraisal of stressors (part of the system of “automatic processes that are activated
in response to stress that are related to but distinct from coping" [Compas, 2006, pg230]), I have gone to pains here to present a case for the involvement of cognitive resources such as academic SE in the regulation of preconscious appraisal processes. Thus, despite the functional separation of automatic and volitional responses, a necessary inductive sequitur of my assertion that SE buffers preconscious stressor appraisals is that other cognitive processes (chief amongst which are coping resources) may also have a significant effect on preconscious systems.

Although there is little direct evidence to support this view, it is well established that people with high SE tend to adopt more adaptive coping styles (Knoll et al., 2005; Terry, 1994), which has led to the suggestion that the protective effect of SE on stress could well be mediated by coping resources (Folkman, 1984; Dweck, 2000). As Ellen Skinner puts it:

"Individuals who have high levels of self-efficacy tend to... cope constructively: regulation is action-oriented and focused on generating strategies, exerting effort, and using outcomes (even failures) as information to shape subsequent strategies" (Aldwin et al., 2011, pg566).

In other words, SE and adaptive coping tend to be strongly correlated (Knoll et al., 2005). As I neither controlled for, nor assessed, coping style in this study, it is entirely plausible that the protective effect of SE identified in this study stemmed, in fact, from individual variations in coping resources via a putative effect on preconscious systems (Putman & Roelofs, 2011). This potential limitation impinges significantly on the construct validity of this study and, therefore, I call for further
investigation in this area to assess whether coping resources, SE, or a combination of both lies behind the distribution of SA-IAT scores reported in this study.

Even if one does not accept the syllogism that, if SE affects preconscious systems so could coping, the premise behind the second and third research questions was that (either directly, or indirectly) stress affects the preconscious appraisal of stressors. The assumption being that stress was experienced homogeneously by the Experimental group as a consequence of exposure to an academic stressor - one or more public examinations. However, because the stress was experienced over a period of weeks, it is inescapable that this will have resulted in activation of "conscious volitional efforts to regulate emotion, cognition, behaviour, physiology" (Compas et al., 2001, pg89) which, in turn, will have instigated efforts to regulate stress and stress-related emotions (Primary Control Coping), efforts to adapt to the presence of the stressor (Secondary Control Coping) and/or efforts to withdraw from stress and stress-related emotions (Disengagement Coping - Connor-Smith et al., 2000). Such coping strategies will inevitably have affected the development of the stress process during the study. Because stressors are known to have multiplicative effects when presented concomitantly (e.g. Lepore & Evans, 1996);

"People cope with multiple environmental stressors, which interact with each other, creating effects which differ from the simple addition of their individual effects" (Martimportugués-Goyenechea & Gomez-Jacinto, 2005, pg867).
Thus it is plausible that the variation in PSS10 data and/or SA-IAT scores between participants owed not to variations in academic SE, but rather to variance in the evaluation of the valence of the PoS stressor resulting from interaction between appraisal processes and pre-activated antecedent-based coping strategies (Mauss et al., 2007). Thus the individual effect of coping resources across the duration of the study could potentially contribute to a Maturation-type error as described by Cook and Campbell (1979). As with my previous point, further research is required to unpick this possibility and thereby establish whether the variance in stress reported here occurred as an effect of academic SE (as I assert), or via a coping-based mechanism.

Aside from the inclusion of history threats, a second limitation germane to the use of quasi-experimental methodologies is the difficulty of demonstrating a causal relationship between experimental variables (Shadish et al., 2002). Although I fully addressed two of Mill's prerequisite 'fundamental conditions' for investigating causal relationships (i.e. a] the cause should precede the effect and b] the cause should be statistically associated with the effect) and took steps to comply with the third 'fundamental condition' (c] there should be no plausible alternative explanations for the effect other than the cause - Rutter, 2007); nevertheless, the statistical measures employed here demonstrated either simple correlations between variables, or confirmed differences between the variance of variables distributed across two discrete populations - neither of which serves to demonstrate a direct causal relationship between those variables. Whilst I have already suggested potential modifications to the methodology (see Chapter 4) which could improve the statistical rigor of this study, I posit that
further work utilising intervention-based procedures would best address this limitation. For example, adopting a false-feedback SE-manipulation protocol such as that employed by Holroyd (Holroyd et al., 1984), or Neufeld and Thomas (1977) would allow for direct assessment of covariance between SE and preconscious stressor appraisal. Similarly, covariance could be assessed utilising the type of SE-dependent stress-reduction training discussed earlier in this Chapter (e.g. Kraag et al., 2006; Masten & Coatsworth, 1998). As the robust exposition of covariance between cause and effect (or the absence of effect when cause is withdrawn) is considered a reliable criterion for the demonstration of causality (Cook et al., 1990), such interventionist methodologies would effectively address this limitation by either establishing "a lawful relationship between cause and effect" (Cook & Campbell, 1979, pg78), or by falsifying (Popper, 1959) my assertion that academic SE protects adolescents from stress via a buffering effect on preconscious appraisal processes. It is my hope that such studies will be conducted expeditiously.

In keeping with the previous point, a third limitation contiguous to the methodology is the lack of comparative baseline data for the participant groups against which to contrast the effect of the PrS tests. In other words, the methodology does not allow for measurement of the effect of pre-testing on performance - an effect frequently described as Testing Error (Shadish et al., 2002). Although this limitation was partially ameliorated through the deliberate use of instruments with well-established normative values, such indirect comparisons did not permit statistical regression of baseline effects within the participant groups. Furthermore, as IATs effects are known to be affected by repeated
testing (e.g. Waters et al., 2010), it is conceivable that the pre-testing process itself obfuscated accurate measurement in the second half of the study. Although pragmatic considerations (e.g. participant numbers) prohibited the use of a Solomon four group design in this study, I recommend that future studies incorporate such methodology to overcome any existing Testing Errors in subsequent research.

**Further limitations to generality.** In addition to the points described above, there are two further limitations to the external validity of my conclusions, both of which relate to variables I controlled in this study. As described earlier, age is intricately linked with the development of academic self-efficacy (Anderman et al., 1999; Urdan & Midgley, 2003), stress (Rudolph & Hammen, 2003; Rutter, 1981) and the ability to cope with stressors (Skinner & Zimmer-Gembeck, 2007; Zimmer-Gembeck & Skinner, 2011). By limiting the study to participants in late adolescence, it is very likely that the conclusions I have drawn about the protective relationship that exists between academic SE, stress and preconscious stressor appraisal are not the same for younger adolescents, children or adults. It is, therefore, important that this study is repeated for participants across those age groups to ensure that the conclusions presented here develop wider external validity.

Similar to the previous point, there are marked differences in the way academic SE develops in boys and girls (Britner & Pajares, 2001; Pajares & Valiante, 2001; see also Hansen, 2009; Pajares, 2002). Similarly, stress is widely known to be gender-specific, both in the response to acute stressors (Matud, 2004) and in the development of chronic stress (Cleary, 1987) and stress-related illness (Weissman & Klerman, 1987).
Furthermore, evidence exists to suggest that the effect of cortisol may differ markedly between men and women (Nielsen et al., 2008) and that women exhibit greater threat-specific attentional biases relative to men (Waters et al., 2007), which implies a direct gender-specific difference at the level of preconscious appraisals. As I concluded for the previous limitation, further research is required to ascertain whether the findings of this study generalise equally to men, perhaps by conducting a repeat of this study in a boys’ school?

Limitations of the instruments. In Chapter 4 I collated a wide body of evidence to support the validity of the PSS10 and ESES instruments. However, despite this evincing material, these tests remain self-reports and are, therefore, subject to the inherent limitations of such tools. For example, as Greenwald and colleagues explain (Greenwald et al., 2002), self-reports are prone to error through the effect of various Response Factors, which include the ability to respond accurately, apprehension related to responding (Rosenberg, 1969) and the problem of response faking (Cronbach, 1990 – see Appendix B for further discussion). As a consequence of these errors, the accuracy, internal validity, and construct validity of these instruments may well have become compromised in this research, which undermines the external validity of my conclusions. To overcome this limitation, confirmatory studies need to be undertaken incorporating physiological measures of stress (e.g. HRV analysis – see Berntson & Cacioppo, 2004) and implicit stressor appraisal (e.g. pupillometry [Laeng et al., 2012] or GSR testing [McGinnies, 1949]), which would serve effectively as triangulation procedures (Johnson & Onwuegbuzie, 2004) to confirm the findings of this study and thereby reinforce the generality of my conclusions. However, the remaining caveat
with this suggestion is that, to date, direct quantitative measures of academic SE have yet to be developed and so, such triangulation procedures would inevitably be forced to incorporate academic SE self-report instruments similar to the ESES utilised in this study. Looking to the future it is my hope that, with continuing advances in neuroimaging, it may soon become possible to map the neocortical system/s and processes that generate academic SE and thus excise the necessity of self-report measures in future studies when direct neurological measures of SE become available.

In addition to the use of self-reports, there are two significant limitations specific to the SA-IAT instrument. Firstly, because the SA-IAT measures the *relative strength* of a pair of associations, it cannot be interpreted as a measure of the *absolute strength* of single associations (Greenwald & Farnham, 2000). I discuss this limitation fully in Appendix B (and, therefore, do not repeat the argument here), but reiterate the conclusion stated there that, because the SA-IAT scores are neither normalised, nor measured according to an absolute interval scale, the data derived from the SA-IAT must be viewed as an ordinal hierarchy *relative to itself*. This greatly limits the degree to which conclusions based on SA-IAT data generalise outside of this study. Secondly, and in keeping with the previous limitation, the SA-IAT instrument has not been tested outside of the environment in which it was conceived. Whilst I have taken (rigorous and, arguably, successful) steps to validate the SA-IAT (see Appendix B) these validation procedures utilised the same students who had originally participated in the initial design of the instrument. Because IAT tools are, at heart, *relative measures*, both the familiarity of stimulus items (Rothermund & Wentura, 2004) and the semantic valence
of those items (Belezza et al., 1986) have a marked effect on the internal validity of IATs (see Hofman et al., 2005; Greenwald et al. 2003; 2009). Although I took steps to control for these effects (see Appendix B), the point is that these steps did not incorporate participants outside of the single school community in which the study was conducted. As I argued earlier in Chapter 4, because individual schools develop a unique personality (Schein, 1985), which varies significantly from school to school (Paterson & Deal, 2009), it follows that the relative associations between stimulus items that underpin the SA-IAT would not hold constant for participants from different schools. Therefore, before assuming that the SA-IAT instrument (and conclusions based on data derived from it), hold validity outside of the initial school community in which the instrument was designed, further work must be conducted to establish the external validity of the SA-IAT instrument, preferably through repetition of the methodology utilised in Appendix B in a number of schools from as diverse a range of cultures and communities as possible.

**Limitations of the statistics.** Aside from arguments pertaining to causality, there are two further statistical limitations that weaken the external validity of this study. Firstly, although I managed to recruit significantly more participants than the minimum target of 52 (see Chapter 4), I did not manage to achieve the recommended target of an Experimental group of 128 participants to cater for a medium effect size (i.e. $\theta \geq .25$ as defined by Cohen [1992]). Accordingly, the respective power of the descriptive statistical measures employed in Chapter 5 lie on a wide range ($.91 \leq \delta \leq .37$) for the effect sizes achieved in this study. As with previously described limitations, repeating the study with a
larger participant cohort would be an appropriate intervention to increase the statistical power of future studies.

The second statistical limitation is that, throughout the study, I have assumed that all recorded variables were homoscedastic - i.e. that the error variance across correlations was constant. However, it is certainly possibly that the relationship between SE and stressor appraisal does become non-linear at extremities, particularly for individuals with significantly higher SE. In this study I did not have enough data to assess this effect directly, but urge future researchers to look at this effect in more detail, perhaps by conducting a large scale regression analysis on participants with high SE (cofactors could include the variables discussed in the earlier part of this section), to analyse the relationship between high SE and stress. The simple addition of a Breusch-Pagan-Godfrey or Goldfeld-Quandt test would serve either to confirm my assumption of homoscedasticity, or highlight a non-linear relationship between SE and stress under some circumstances.

Summary

At the very beginning of this project I explained that “one of the guiding motivations behind this work is a desire to contribute, albeit in a very small way, towards the development of a national curriculum that incorporates a focus on stress resistance, or anti-stress training.” The conclusions of this study unerringly point towards SE as a crucial founding stanchion of such a programme, because of the protective role SE so clearly plays in shielding adolescents from stress. This assertion is doubly reinforced by the success of class-based schemes (Lohaus, 2011) and
outdoor-based activities (Masten & Coatsworth, 1998), which have successfully used SE-development training to reduce stress. It is my sincere hope that, if this study achieves any tangible outcome, it will be the expansion of such programmes of study in the UK over the next few years.

In this study I have attempted to explain the protective effect of SE from a dual process perspective of stress (Compas, 2004), focusing specifically on the role of preconscious appraisals in the genesis of acute stress. This angle has not been explored previously in stress research, so it is exciting that this study has not only highlighted the important role implicit appraisals can play in determining the magnitude of the acute stress process; but also that SE may exert all or part of its protective effect by decreasing the degree to which stressors are implicitly appraised as stressful. Although I have been unable to identify whether this effect occurs at the preconscious level, or via a top-down influence; nevertheless, it opens the door for future research to establish the exact mechanism by which SE modulates or buffers preconscious stressor appraisals and thereby SE protects adolescents from acute stress.
Chapter 7: Implications for Practitioners

As discussed in Chapter 3, a significant body of research indicates that SE-enhancing curriculum programs such as that used by De Wolf & Saunders (1995) and Hampel (Hampel et al., 2008) may offer effective (Kraag et al., 2006) and long-lasting (Masten & Coatsworth, 1998) interventions for reducing the academic stress frequently experienced by adolescents at school (Lohaus, 1990; Currie et al., 2004). In this chapter I briefly summarise recent progress in this field and make suggestions which may help teachers and educationalists to reflect on their own pedagogy with a view to increasing the academic SE of the young people within their charge.

Citing evidence linking academic underachievement, violent behaviour and physical illness (Altshuler & Ruble, 1989; Compas & Hammen, 1994) with maladaptive coping strategies (e.g. Olbrich, 1990), Segal’s (1983) pioneering stress-reducing intervention program established conclusively that teaching school children adaptive coping strategies reduced not only their level of self-reported stress, but also increased their long-term academic achievement as well. This effect has since been replicated by other studies (e.g. Forman, 1993; King et al., 1998) which have specifically highlighted interventions that boost either children’s emotional coping skills, or their problem-focused coping strategies (e.g. De Wolf & Saunders, 1995; Dubow et al., 1993) as being particularly effective in reducing academic stress. Although such programs most likely mediate the bulk of their effect through a coping-based pathway, authors such as Gerda Kraag have been quick to identify an underlying (and relatively
ubiquitous) secondary effect of these programs – namely, a significant accompanying increase in SE (Kraag et al., 1996).

This observation led to the advancement of a different group of stress intervention curricula – those that aim to reduce the development of stress, rather than increasing children’s ability to cope with the stress once it has already been experienced. Such programs can be typified by that of Schinke, Schilling and Snow (1997), who used a combination of communication training, education in self-instruction and lessons in problem-solving skills to raise the self-reported SE of 278 6th grade children, resulting in significant decreases in stress and anxiety and long-term heightening of self-esteem, interpersonal assertiveness and a mild (though statistically significant) elevation in academic achievement. Other, similar school-based training schemes involving social problem solving (Elias et al., 1986; Weissberg et al., 1981), social-emotional awareness (Klingman & Hochdorf, 1993) and relaxation (Zaichkowsky & Zaichkowsky, 1984) have also reported equivalent increases in SE and a resultant reduction in perceived stress, measured by a variety of direct and indirect measures.

Although these SE-based curricula all utilise specific training interventions, I posit that the core themes of these curricula generalise specifically towards three of Bandura’s (1997) four sources of SE, namely Enactive Mastery, Vicarious Experience and Verbal Persuasion. Apropos, by explicitly and purposively equipping young people with strategies either for coping with, (Dubow et al., 1993) or devising solutions for common

---

65 The fourth source is affective state. However, as SE plays a key role in determining affective state, it is likely that SE training indirectly benefits this area as well.
academic stressors (Weissberg et al., 1981). I argue that such intervention programs effectively bolster SE via repeated, reinforced and positive experience of Enactive Mastery. Additionally, by doing so in a carefully controlled peer group setting, students also gain considerably by Vicarious Experience gained through viewing their peers learning to tackle stressors. For example, in a large study across a number of Chicago schools, Yasutake and Bryan (1996) found that, in peer settings, students;

"Had a profound influence on one another... [and that] the combination of peer tutoring with attribution training influenced self-perceived competence in a dramatic and positive way" (Yasutake & Bryan, 1996, pg23)

Similarly, the combination of carefully selected and demand-specific group tasks, also serves to emphasise the role students serve in reinforcing the collective development of SE. Thus, as emphasised by Fencl and Scheel: “the type and nature of teaching strategies used in the classroom can and do make a difference to students' self-efficacy” (Fencl & Scheel, 2005, pg 21 - original emphasis) and, therefore, in their subsequent susceptibility to stress. Indeed, extensive research has repeatedly confirmed this (e.g. Bandura & Schunk, 1981; Dweck, 1975; Schunk, 1985; Wood & Locke, 1987), leading to the conclusion that, with minimal training and effort;

"Teachers can modify their instructional strategies... and this can result in increases in their students' self-efficacy and resistance to stress" (Siegle & McCoach, 2007, pg279).

The next logical question, then, is how can teachers and educators modify their pedagogy to facilitate this?
A large body of research has identified teacher feedback as playing a particularly significant role in the development of adolescents' SE (Schunk, 1984; 1985). Not only is the timing of feedback believed to be important (i.e. intervention is most successful when given within a few hours of the learning activity - Schunk, 1984), but also the style and content of feedback has been found to have a substantial effect on the development of SE (Schunk, 1989a; 1989b). For example, feedback that reinforces recent success (Bandura, 1997) has a markedly positive impact on SE and academic performance. Similarly, feedback that develops this to emphasise the balance between ability and effort has been reported as most effective in developing SE (Schunk, 1989b), particularly when poor performances are viewed in terms of lack-of-effort explanations and successes in terms of ability (Siegle & McCoach, 2007). This effect seems to be particularly helpful in overcoming gender-specific differences in the development of academic SE (Kenney-Benson et al., 2006). Furthermore, feedback is reported to be enhanced when given systematically - i.e. to all learners en bloc, rather than in an unsolicited fashion (Zimmerman & Martinez-Pons, 1990), possibly because students tend to associate spontaneous help with teachers targeting underachieving pupils (Graham & Barker, 1990), which is erosive to SE.

Similar to effective teacher feedback, structured goal-setting can have a substantial effect on improving SE (Bandura, 1986; Schunk, 1989b - see also Appendix A): as "when children can easily gauge their progress against a goal, their perception of improvement enhances their self-efficacy" (Schunk, 1989a, pg17). Goals that are unambiguous (Dweck & Leggett, 1988) and relate to specific performance standards tend to be
the most effective (Gutman, 2006). Equally, goals that let students decide for themselves how to break up larger goals into smaller, attainable ones tend to facilitate better the development of SE. Finally, goals are most effectual when students perceive them to be difficult, but achievable (Siegle & McCoach, 2007)

Reiterating the advice already given, Margolis and MaCabe (2006) assert that practitioners should; 1) plan moderately challenging tasks, 2) teach specific learning strategies, 3) capitalise on student choice and interest, and 4) reinforce effort and correct strategy use. However, they also emphasise the role of peer modelling in cultivating academic SE as this serves both to reinforce vicarious experience and also verbal persuasion, particularly if the peer model has more competence in the areas of focus than the targeted pupil - a pairing strategy Schunk (1989a) found to be more effective even than teacher-based modelling. Developing this area further, Siegle and McCoach (2007) advocate teachers employing “a variety of ways to use models” (pg286), including videotaping, peer tutoring, work groups and class demonstrations, all of which “can help to exploit the power of models” Siegle and McCoach (2007, pg286). Furthermore, early research has established that self-modelling techniques (e.g. where a videotape of a student performing a desired behaviour is made, with any undesired or unsuccessful behaviours edited out) also constitute promising tools for fostering the development of SE in young adults (Bray & Kehle, 2001).

Complementing the guidance already presented, Demos and Foshay (2010) advocate strategies for extending and expounding classroom motivation. Not only does this pedagogy seem to foster academic engagement
particularly effectively (which, in turn is linked to higher grades, higher test scores, lower drop out rates [Siegle & McCoach, 2005], and correlates well with achievement and persistence [Klem & Connell, 2007]), but it has also been reported to lead directly to the promotion of academic SE (Bandura, 1995b). Although numerous strategies for increasing engagement and motivation have been suggested (e.g. see Coiro & Dobler, 2007), Guthrie and Davis (2003) identify three categories of intervention that have proven particularly effective; namely 1) ‘pegging’ or linking tasks and activities with children’s underlying interests, 2) the development of strategic goals that expand the previous point in tandem with the learning objectives, and 3) schemes that enhance literacy and/or lead to reading strategy development. Of these, the latter is particularly pertinent as disengagement is frequently connected with reading difficulty:

"Given the importance of reading to school achievement, limited success with reading is likely connected to poor grades, difficulty with curricular demands, and pervasive disenfranchisement with school requirements" (Demos & Foshay, 2010, pg58).

Although a full review of techniques to foster the development of engagement through literacy is beyond the scope of this chapter, I point the reader towards literature from the Centre for Applied Special Technology (CAST, 2008), whose research into universal learning guidelines has made significant progress within this field.

One final area for consideration is the adoption of heutagogical curricula, which have been found to correlate reliably with increases in SE
(Bandura, 1997). For example, Alfassi (2003) found students enrolled in learner-centred study programs consistently out-perform control groups engaged in existing, more pedagogical curricula. His research has been replicated by others, who conclude that:

“Schools have an opportunity to build self-efficacy through instructional design methods that focus on a learner-centred approach” (Daumer, 2006, pg 32).

This conclusion mirrors the material I presented earlier documenting the success of outdoor training curricula programs (Thurber et al., 2007; McKenzie, 2000; Masten & Coatsworth, 1998) which, by necessity, have centred on the development of specific heutagogical skills. Equally, parallels can be drawn with coping-based schemes that have successfully enhanced adolescents’ SE through specific aptitude training (Hampel et al., 2008). Though wide-scale, I suggest that by emulating the ethos of these large-scale stress intervention programs there is much that can be done on a small scale to make learning, learning environments and pedagogy more pupil-centred. This, in conjunction with the recommendations made by researchers such as Margolis and MaCabe (2006), may well enable teachers and educators to foster more easily the development of academic self-efficacy in young people, and thereby help protect our children from many of the maladaptive effects of academic stressor exposure.
References:


Bargh, J., Chaiken, S., Raymond, P. & Hymes, C. 1996. The automatic evaluation effect: Unconditional automatic attitude activation with a


Barrett, L. 2009. Variety is the spice of life: A psychological construction


Cohen, S. 2012. PSS Website. [www.psy.cmu.edu/~scohen/scales.html](http://www.psy.cmu.edu/~scohen/scales.html) (accessed 03.11.12).


Coiro, J. & Dobler, E. 2007. Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet, Reading Research Quarterly 42(2): 214 - 257.


Theory, Methods, Research, New York: Oxford University Press, 141 - 156.


Ilg, R., Wohlschlager, A., Gaser, C., Liebau, Y., Dauner, R. & Woller, A. 2008. Gray matter increase induced by practice correlates with task-


Jensen, A. 1993. Why is Reaction Time Correlated with Psychometric g? 
*Current Directions in Psychological Science* 2: 53 - 56.


Kappas, A. 2006. Appraisals are direct, immediate, intuitive, and unwitting... and some are reflective..., *Cognition and Emotion* 20(7): 952 - 975.


vaccine in older adults, *Proceedings from the National Academy of Science* 93: 3043 - 3047.


Nosek, B. & Smyth, F. 2005. *Implicit and explicit attitudes are related, but distinct constructs*, University of Virginia: Charlottesville, VA.


between psychological stress and telomere length, *Brain, Behaviour and Immunity* (in press).


Öhman, A. & Soares, J. 1998. Emotional conditioning to masked stimuli: Expectancies for aversive outcomes following non-recognized fear-


Schwarzer, R. 2011. Everything you wanted to know about the General Self-Efficacy Scale but were afraid to ask. www.userpage.fu-berlin.de/~health/faq_gse.pdf (accessed 17.11.11)


prevention programmes: NIDA Research Monograph 170, National Institute on Drug Abuse, 147 - 164.


Solberg, V. & Villarreal, P. 1997. Examination of self-efficacy, social support, and stress as predictors of psychological and physical distress


& C. Judd (Eds.) *Handbook of research methods in social and personality psychology*, New York: Cambridge University Press, 40 - 84.


Appendix A

The effect of SE on academic performance
The effect of SE on academic performance

High SE has been shown to predict performance in a variety of different assessment protocols (De Vries et al., 1988) including musical recitals (McPherson & McCormick, 2006), reading tasks (Schunk & Rice, 1993), athletics (Zimmerman & Kitsantas, 1996) and examinations (Lane & Lane, 2001) across a wide range of academic subjects (see Bandura, 1997 for a review). Meta-analyses conducted by Stajkovic and Luthans (1998) and Multon and her colleagues (Multon et al., 1991) found that self-efficacy accounted for, respectively, 28% of the improvement in work-related performance and 14% of variance in academic accomplishment, indicating that “efficacy beliefs contribute significantly... to performance accomplishments” (Bandura, 2001a - pg4). Furthermore, growing evidence exists demonstrating that SE self-reports are strongly predictive of performance (Bandura, 1997), more so even than academic ability (Jinks & Morgan, 1997; Bandura, 1993). For example, during research for his thesis, Collins (1982, cited in Zimmerman & Cleary, 2006) divided high-school mathematics students into high-scoring and low-scoring groups according to their perceived maths efficacy. He found that, irrespective of their underlying ability at maths, students in the high efficacy group “discarded faulty solution strategies more quickly, reworked more failed problems, and achieved higher maths performance” (Zimmerman & Cleary, 2006, pg56). The same effect was also witnessed by Schunk and Gunn (1986, pg243), who found that “SE makes an independent contribution to academic performance” in a division-based maths task, which could not be attributed to performance attributes, task strategy or to effort. Their work was repeated and the findings further corroborated by Bouffard-
Bouchard and her colleagues (Bouffard-Bouchard, Parent & Larivee, 1991), who also concluded that SE exerts a significant influence on performance.

Similar supportive evidence demonstrating a link between SE and performance can be drawn from experiments examining the opposite procedure to Schunk and Gunn (1986), where poorly calibrating SE with the difficulty of a problem-solving task causes diminished performance, which is, again, not determined by students' academic capability (Bandura & Wood, 1989). However, the most compelling evidence strengthening the association between SE and academic performance comes from prospective studies where existing SE is built up through training interventions and its effect on performance is then observed. The outcome of the training reliably produces a corresponding positive effect on accomplishment and, as with studies cited previously, the effect supersedes underlying academic ability. For example, in a word-selection task, Bouffard-Bouchard (1990) demonstrated that students whose efficacy had been raised through false positive feedback not only achieved a higher solution-rate than the control group, but also employed better strategic flexibility and had higher levels of aspiration. The same result was observed by Cervone and Peake (1986) for an anagram and graphical problem-solving task. A large body of work documenting the affect of intervention-based training programs on students' SE and performance has been carried out by Dale Schunk (e.g. Schunk 1981; 1987). “In these investigations Schunk and his colleagues not only demonstrated the sensitivity of efficacy beliefs to instructional interventions, but also the mediational role of these beliefs in explaining changes in learners’ self-regulation and achievement outcomes” (Zimmerman, 2000b, pg89). Similarly, research investigating Zimmerman's
model of self-regulation (Perels, Gurtler & Schmitz, 2005) also underscores the strong link between SE and performance, establishing SE “as a potent mediator of students' learning... [which] confirms the historic wisdom of educators that students’ self-beliefs about academic capabilities do play an essential role in their motivation to achieve” (Zimmerman, 2000b, pg89).

However, the exact nature of this ‘essential role’ has been extensively debated. Bandura (1982, pg126) argues for a “causal contribution of self-efficacy to action”: in other words, SE itself has a causative affect on behaviour. However, this position has come under some criticism. Lee (1989) dismisses SE as being non-scientific - SE is “a vague descriptive model, not an explanatory theory” (Lee, 1989, pg121, my emphasis). Hawkins (1992, pg252; see also 1995) explains Lee's point clearly: he argues that, as SE is “a hypothetical construct which is invoked to help understand a pattern of observed behaviour” it cannot be a de facto independent variable - “it [SE] does not exist” (Hawkins, 1992, pg255) Consequentially, it is inappropriate to analyse behavioural observations with an a priori concept of SE as this not only makes SE effectively into an unfalsifiable concept (Lee, 1989) but also constitutes circular logic. To illustrate the point, Hawkins draws the example of a study (Bandura et al., 1980) frequently cited by Bandura (e.g. Bandura, 1982) as evidence of causality between SE and behaviour. In this quasi-experiment Bandura and his colleagues used a behaviour-training program to increase phobic participants’ SE from a low starting level to a much higher degree. Measuring performance by tolerance to phobia exposure Bandura noted that, not only did performance improve significantly after training, but that the increase varied in proportion to the participants' perceived SE.
He concluded that SE had directly led to the increase in tolerance. However, applying the principle of Occam’s Razor, Hawkins argues that the most parsimonious explanation is that the behavioural training program itself affected the participants’ behaviour and, therefore, that the behavioural training, rather than the improvement in SE, must be viewed as the independent variable in this study:

“A hypothesis involving a theoretical term not defined independently of the behaviour giving rise to it is always less satisfactory than a functional account which involves no special concept at all” (Hawkins, 1992, pg256).

Hawkins (1995) argues that, in this example, SE simply acts as a record of experience and is, therefore, not itself causally linked to performance.

Bandura’s polemic response (Bandura, 1995a) effectively dispatched Hawkins’ ‘Skinnerian functionalism’ on the grounds that:

“Evidence from countless studies demonstrate[s] that perceived self-efficacy contributes independently to subsequent performance when variations in prior performance are controlled” (Bandura, 1995a, pg180).

This evidence is extensive and compelling (e.g. see Bandura, 1984; 1997) and is further corroborated by studies examining the synergistic relationship that exists between SE and performance in some work settings (Gist & Mitchell, 1992). In such environments the ‘malleability’ of SE is strongly determined by performance outcomes, such that self-reinforcing ‘efficacy-performance spirals’ are generated, both in individuals and within groups and organisations (Lindsley, Brass & Thomas, 1995). These ‘spirals’ constitute transparent evidence of a causal link
between SE and performance and, whilst other (more valid) criticisms of SE will be addressed later on, for the remainder of this piece I take the position that “students’ self-efficacy beliefs... play a causal role in students' development and use of academic competencies” (Zimmerman, 2000b, pg89).

How SE affects performance outcomes

1. Increased Persistence. Measured under various guises (e.g. 'number of attempts' [Schunk, 1982] or 'effort' [Salomon, 1984; Schunk, 1983b]) children with higher levels of SE reliably demonstrate greater task persistence. For example, when given an insolvable wire- or word-puzzle, Zimmerman and Ringle (1981) found that children with higher self-reported SE spend longer attempting to solve the puzzles before giving up. The same effect was observed by Zimmerman and his colleagues for college students sitting exams (Zimmerman, Bandura & Martinez-Pons, 1992) and is clearly evident in meta-analyses where, for example, Multon and her colleagues (Multon et al., 1991) found that SE was independently responsible for 12% of persistence. SE is also heavily implicated in engagement theory (see Linnenbrick & Pintrich, 2003 for a review) where students with high SE have been shown not only to display longer periods of behavioural engagement, but also reliably demonstrate deeper processing strategy and better cognitive engagement as well (Pintrich & Schrauben, 1992), sometimes by up to 10% more than students with lower SE (Pintrich, 1999). Further evidence linking SE with persistence can be drawn from the work of Paul Pintrich on the positive impact of motivation on learning (see Pintrich, 2000b for a review). Although a separate domain to persistence, motivation is strongly predictive of effort (both
behaviourally and cognitively – see Linnenbrick & Pintrich, 2003) and has been shown repeatedly to correlate positively with SE (Pintrich & Schunk, 1996). Pintrich (2003) conceptualises motivation as a multi-factorial construct that over-laps considerably with interest (see next section), emotion, affect and SE. Although he acknowledges that it is very difficult to ascertain the relative causal ordering of SE and motivation - i.e. does SE cause motivation, or vice versa? - in this case, regardless of whether the tail wags the dog or the dog wags its tail, his conclusions very much establish a central role for SE in regulating student performance via effects on persistence and motivation.

2. Increased Interest. In his excellent book Schunk (1989) presents evidence suggesting that students with high levels of specific academic SE develop a greater depth of interest in that domain. This has subsequently been confirmed by others (e.g. Bandura, 1997), who found that, independent of ability, personality or outcome expectancy (Donnay & Borgen, 1999) SE may contribute up to 27% of variance in interest (Lent, Brown & Hackett, 1994; see also Rottinghaus et al. [2003], who concluded from their [much larger] meta-analysis that SE makes a lower, but nevertheless significant, contribution to interest). However, research by Silva (2003) suggests that SE may exert a more complex, 'quadratic influence' on interest - i.e. initial increases in SE have a marked positive effect on interest, which gradually diminishes until the task reaches a point where "success seems completely certain and the task is thus uninteresting" (Silva, 2003, pg237 - abstract). This effect was first posited by Bandura (1986b), who also postulated the existence of a low-SE 'threshold' before an increase in interest can be observed. Whilst evidence exists to support the second proposition (Lenox & Subich, 1994),
the exact nature of the link between SE and interest requires further investigation as, although the quadratic relationship documented by Silva has been observed in some studies (Tracey, 2002), it has not been replicated in others (Lenox & Subich, 1994).

3. **Goal-Setting.** A wealth of literature demonstrates not only the direct and positive effect of goal-setting on task performance (see Locke et al., 1981; Locke & Latham, 1990 for reviews), but also the significant influence on accomplishment SE contributes via its mediating effect on the goal-setting process (Locke, Frederick & Bobko, 1984). For example, in a study of male varsity basketball players Cleary and Zimmerman (2001) found that, although ‘expert’ sportsmen tend to set approximately the same number of goals compared to ‘non-experts’, there was a pronounced disparity in the type of goal they selected, and that the degree of variation in goal type was very closely correlated with the athletes’ SE. Cleary and Zimmerman (2001) concluded that high SE leads people to set more specific goals: an observation that has been confirmed by many others (see Locke & Latham, 1990). Furthermore, in addition to more explicit goal-setting, SE also causes people to adopt goals that are both more realistic (McClelland, 1985) and more challenging (Schunk, 1990), which has the combined effect of increasing the commitment students are prepared to make to achieve the goals (Zimmerman, 1995; see also Point 1 above).

As well as goal-perseverance and goal-type, people with high SE make two further enhancements to their goal-setting process: a) they adopt a hierarchical goal structure that includes ‘paced’ or ‘proximal’ shorter-term goals (Bandura’s 1986b), which has the same positive effect on
performance outcome as changing goal-type (Bandura, 1997); and b) they implement different strategies to achieve their goals, which include changes to the performance monitoring and management process, a more specific performance-goal evaluation mechanism and a greater degree of self-evaluation with respect to performance (Zimmerman, 1989). Therefore, through its constructive and synergistic mediation of multiple steps in the goal-setting process, SE has a particularly potent and positive effect on achievement. However, although Bandura (1997) asserts that the effect of SE on goal-setting is linear - i.e. goal-setting has no reciprocal effect on SE ("people do not choose the goal of swimming a treacherous body of water and then wonder whether they have the swimming capabilities to reach the opposite shore" [Bandura & Cervone 1986, pg104]), this view has been increasingly undermined by studies observing a bidirectional relationship between SE and goal-setting (Garland, 1985). Schunk (2003) postulates that "specific goals... promote efficacy because it is relatively easy to evaluate progress toward an explicit goal" (Schunk, 2003, pg163-164); therefore, it now seems likely that, in addition to augmenting performance, SE itself is also increased indirectly through the processes outlined above.

4. Self-Regulation Strategy. Early investigation into metacognitive learning processes - processes that develop "awareness of and knowledge about one's own thinking" (Zimmerman, 2002, pg65) - clearly demonstrated that students' use of self-reflective learning strategy is well correlated with academic success (Zimmerman & Martinez-Pons, 1988) and is highly predictive of academic achievement (Zimmerman & Martinez-Pons, 1986) independently of prior attainment (McClelland, Morrison & Holmes, 2000) and of ethnicity, gender and other personal
differences (Yen, Konold & McDermott, 2004). This is particularly the case for students of lower ability (Zimmerman & Martinez-Pons, 1986). These initial findings gave rise to a very large body of research on the "self-directive process by which learners transform their mental abilities into academic skills" (Zimmerman, 2000a, pg18), which quickly became known as **Self-Regulated Learning (SRL)**.

Many different models of SRL have been proposed (see Zimmerman & Cleary, 2001; Boekaerts et al., 2000 for reviews) and, although some (e.g. McCombs & Marzano, 1990) propose a general agentic 'self-system' that mediates between personal and contextual characteristics and performance (Pintrich, 2004), the majority of models identify separate organisational components within SRL, amongst which SE has been implicated as particularly significant (Zimmerman, 1990). Consequentially, SE features in the majority of models of SRL (Zimmerman & Cleary, 2001). For example, in Zimmerman's model of Self-Regulation (see Figure A1 below), SRL is conceptualised as a cyclical, relational process consisting of three discrete phases. The 'Forethought Phase' is perceived as a reinforcing transaction between task-analytic processes (planning and goal setting) and self-motivation beliefs (in which SE plays a pivotal role), which collectively mediate the 'Performance Phase' of SLR.
Zimmerman presents compelling evidence to support the influential role of SE in his model of SRL. For example, in a dart-throwing task Zimmerman and Kitsantas (1997) found that girls who used SE to inform their goal-setting strategy not only achieved the best performance, but also developed substantially higher SE than girls who used different task analytical processes. The same effect was observed in a word-sorting task (Zimmerman & Kitsantas, 1999). In the opposite scenario Pintrich (2000a) conducted a longitudinal study into the effect of different types of forethought strategies on performance in maths. He found that students who were coached to use outcome-specific targets only (i.e. ones that made no use of SE) exhibited significant decreases in SE, increased negative affect and were more likely to withdraw effort and engagement in lessons. From this study (and others) Pintrich (2003) drew two important conclusions: 1) SE heavily influences motivation (see earlier) which is in keeping with Schunk and Swartz's (1993) observation that persistence with learning strategy is positively correlated with SE (Schunk & Swartz, 1993); and, more importantly, 2) that SE leads
students to adopt different learning strategies, chief amongst which are those that focus on planning, systems of self-monitoring and time-management (Pintrich, 1995). This was Zimmerman's thinking as well (Zimmerman & Martinez-Pons, 1992) and, although different SRL theories each invoke SE in subtly different ways, it is highly likely that much of the link between SE and performance occurs indirectly through SE-induced mediation of SRL strategy.

5. **Anxiety.** “When people see themselves as ill-equipped to cope with potentially threatening events, they will become anxious” (Muris, 2002, pg338). In other words, students suffer from achievement anxiety because they have a low perceived SE to manage academic demands (Meece et al., 1990; Pintrich & De Groot, 1990). Anxiety has a pronounced negative effect on performance as demonstrated in correlation path-analyses (Bandalos et al., 1995; Williams, 1991) and in experimental induced-stress studies, where test-anxious students make appreciably more performance errors than control groups (Kurosawa & Harackiewicz, 1995). This is in keeping with students’ own post-examination reports, which reliably ascribe performance deficits to high levels of test anxiety (Cassady & Johnson, 2002). Explanations for this effect have suggested that the anxiety-state may obfuscate memory retrieval cues, either via cognitive interference from other thoughts (Schwarzer & Jerusalem, 1992), physiological stress (Sarason, 1984) or by anxiety-related disruption of pre-test conceptualisation, processing or encoding processes (Naveh-Benjamin, 1991). However, given Khul’s (1982) finding that goal-setting is also undermined by anxiety and other studies demonstrating anxiety-generated interference on other performance-related factors (e.g. anxiety contributing to decreased perseverance -
see Cassady & Johnson [2002] for work on anxiety-related procrastination) it is likely that the full impact of anxiety on performance is multi-factorial, complex and of considerable significance, particularly when the perception of threat is significant enough for anxiety to become stress.
Appendix B

Institution Focused Study
Developing, Testing and Validating the Stressor Appraisal Implicit Association Test (SA-IAT): an Institution-Focused Study

Introduction: why the IAT?

Procedures measuring implicit cognitive processes have proliferated extensively since Greenwald and Banaji (1995) first coined the term ‘implicit social cognition’. For example, in 2005 Nosek and colleagues (Nosek et al., 2005) recorded that, between 1998 and 2004 more than 120 papers had been published using the most popular implicit methodology – the Implicit Association Test (IAT; see Greenwald et al., 1998). In 2011 that figure was close to 3000, which accounted for only 43% of studies utilising implicit methodologies. The grand total stood at 6,282 and included a range over twenty different implicit measurement procedures in what Nosek (Nosek et al., 2011, pg152) termed “a veritable cottage industry... for measuring implicit social cognition.” Such implicit procedures are now also being applied to a wide range of fields across psychology and related disciplines, ranging from emotion regulation (Mauss et al., 2006; Chen & Bargh, 1997) to addiction (Wiers & Stacey, 2006), or from religious views and political affiliation (Nosek et al., 2007) to sexual attraction (Gray & Snowden, 1998).

Traditionally, self-report instruments are employed to measure cognitive processes. However, in the assessment of implicit cognition self-report methodologies are notoriously inaccurate, suffering from the twin limitations of Introspective Limits and Response Factors (Greenwald et al., 2002). Briefly, Introspective Limits describe the ability of participants to report on the intended content domain. As implicit
cognition occurs outside of subjective awareness (Epstein, 1994; Nisbett & Wilson, 1977) patients are unable to self-score accurately because they lack explicit awareness of the content domain assessed by the instrument (Greenwald et al., 2002). This leads to the second limitation, whereby the participant’s willingness to report about him- or herself becomes compromised by various Response Factors, which include the ability to respond accurately, apprehension related to responding (Rosenberg, 1969) and the problem of response faking (Cronbach, 1990). Taken together, these reservations severely limit the use of self-report methodologies in implicit cognitive research and are the main reason for the proliferation of implicit techniques such as the IAT.

As stated earlier, of the numerous implicit methodologies available, by far the most popular is the IAT, which accounts for just under half of all studies in implicit cognitive research (Nosek et al., 2011). No doubt part of the success of the IAT is owed to its ‘ease of implementation’ (Greenwald & Nosek, 2001) and the large effect sizes it produces (Greenwald & Nosek, 2001). Similarly, because the IAT is open source and free to use it is, understandably, appealing to thrifty researchers and insolvent teachers conducting research degrees. However, the real allure of the IAT is its ‘versatility and adaptability’ (Meade, 2012), which has led not only to its wide and varied application, but also generated sufficient data to assess rigorously the validity of the instrument (for a review see Greenwald et al., 2009; see also later section on validity of the IAT). Although alternative methodologies investigating implicit association do exist (e.g. the ‘Go / No-Go’ Association Task [e.g. Nosek 

---

66 This is often subsumed into the wider category of the demand characteristics of responding (Orne, 1962)
Banaji, 2001], the Lexical Decision Task [Wittenbrink et al., 1997] or Stroop-type procedures [e.g. Cox et al., 2006]), these methodologies are considerably less versatile than the IAT and, therefore, tend to be employed in research less frequently than the IAT. Consequentially, the validity of these methodologies is less well documented. This is because: a) there has been less call (or funding) for studies establishing the validity of these methodologies and, b) currently, insufficient data exists to establish test validity based on meta-analytic research. Therefore, of the lexical methodologies, the IAT is the obvious preferred choice.

However, a variety of non-lexical implicit techniques are also available, including a number of affective priming processes (e.g. Fazio et al., 1986; De Houwer & Eelen, 1998; Winkielman et al., 1997) or implicit activity-based procedures, such as the Approach-Avoid Task (Rinck & Becker, 2007) or the Stimulus Response Compatibility Task (Mogg et al., 2003). Similarly, a number of physiological techniques have been developed that indirectly correlate with implicit cognition (e.g. pupillometry [Laeng et al., 2012] and GSR tests [McGinnies, 1949]). However, all of these IAT-alternatives have in common three significant limitations: 1) they require the use of expensive laboratory equipment, 2) they are designed for (and are generally employed in) experiment-type methodologies, which compromises their use for in situ research in a school environment, 3) by nature of their non-lexical design, they are less applicable to the kind of academic stressors found in such environments. Therefore, for the research methodology put forward in Chapter 4, the IAT constitutes the preferred instrument for measuring implicit cognition.

How Does the IAT work?
In essence, the IAT is “a general-purpose procedure for measuring strengths of automatic associations between concepts” (Greenwald et al., 1998, pg1465). The key underlying concept is the assumption that it ought to be easier to make the same behavioural response (a key press) to concepts that are strongly associated, than to concepts that are weakly associated (Nosek, Greenwald & Banaji, 2005). The IAT is, therefore, a timed classification test in which participants use two response keys to sort items into opposing groups. Each group comprises two oppositional sub-components: the initial target categories (e.g. 'exams' and 'holidays') and the associated concept attributes (e.g. 'stressed' and 'relaxed'). These target and attribute categories are combined in two different ways - i.e. the combination of 'exams' + 'stressed' press left and 'relaxed' + 'holidays' press right, is compared with the combination 'exams' + 'relaxed' press left and 'holidays' + 'stressed' press right. The IAT effect is the difference in reaction times between these two sorting conditions, based on the idea that, when two concepts are associated, sorting is easier (and, therefore, will occur faster and with fewer errors)\(^{67}\).

The exact nature of the implicit cognitive and preconscious mechanism/s that underpin the IAT effect is the subject of some considerable debate (see Fazio & Olson, 2003 for a review). However, a key study by De Houwer (De Houwer et al., 2001; see also De Houwer & Hermas, 2001) strongly suggests that the effect owes to a preconscious stimulus categorisation process, rather than an automatic, valence-determined reflex. In a series of important experiments De Houwer asked participants to categorise famous people into 'British'/'foreign' target

\(^{67}\) For worked examples of the IAT, see www.implicit.harvard.edu.
categories, which were paired with 'positive'/‘negative’ attributes. The list of famous people was deliberately structured to include both positively- and negatively-valenced British and non-British people (e.g. Princess Diana, Albert Einstein, Adolf Hitler etc) which, De Houwer hypothesized, ought to influence the categorization process if stimulus valence is important in generating the IAT effect - i.e. IAT scores would produce markedly different results depending whether 'British' accompanied 'positive' or 'negative' attribution characteristics. This is not what he found, suggesting that the specific valence associated with the famous individual names did little to affect the IAT test. This conclusion is also supported by Mitchell and his colleagues (Mitchell et al., 1999), who found that participants produced different IAT scores when asked to categorize the same list of famous white or black politicians and athletes into either 'white'/‘black' or 'politician'/‘athlete' categories - i.e. the stimulus itself was largely irrelevant: the IAT effect was produced by the specific evaluation of the stimulus in terms of its association with the target categories. In other words, the IAT effect is generated by the processes that appraise the stimulus, rather than any specific property of the stimulus itself (Banaji, 2001)68.

It might be tempting to assume that an Implicit Association Test measures only implicit, or preconscious, appraisals as some have suggested (Cunningham et al., 2001). However, this view is limited for two reasons; 1) as I have demonstrated in Chapter 2, both conscious and preconscious systems are capable both of self-regulation and also of regulating each

68 For a review of appraisal processes, please see Chapter 2.
other: in effect, they are inseparable and interdependent (Sherman et al., 2008$^{69}$), 2) as Conrey and her colleagues highlight:

"No task is 'process pure.' It is technically impossible that any task that requires observable responses depends entirely on automatic processes and not at all on controlled processes. Moreover, it is quite unlikely that any task depends entirely on controlled processes and not at all on automatic processes. Rather, most, if not all, of the behaviours researchers wish to understand will be influenced by simultaneously occurring automatic and controlled processes that influence one another" (Conrey et al., 2005, pg470).

Thus, because the IAT depends on behavioural responses, it must be considered as a measure of the combined influence of preconscious appraisal processes and their regulation by cognitive systems.

This view is well supported by neuroimaging studies, which have repeatedly shown that participants undertaking IATs show increased activity in subcortical stimulus appraisal systems (Cunningham et al., 2003), particularly the amygdala (Phelps et al., 2000), but only if stimuli are presented for 25ms or less (Cunningham et al., 2004). Beyond this window, activity also registers in neocortical areas (Chee et al., 2000), particularly the PFC and ACC (Aron et al., 2004), which are heavily involved in cognitive appraisal processes (see Chapter 4). Furthermore, other studies (Richeson et al., 2003; Cunningham et al., 2004) have found a clear correlation between IAT scores and PFC activity, leading to the conclusion that, whilst preconscious appraisal processes mediate the

$^{69}$ This has led to the movement away from dual processes models of appraisal towards diffusion-based models such as the Quad Process model (see Beer et al., 2008 for a review).
initial IAT response\textsuperscript{70}, the activity is carefully regulated by overlying cortical processes (Stanley et al., 2008). Thus, according to Chee and colleagues (Chee et al., 2000), the IAT effect is generated by cortical processes over-riding preconscious responses. According to this view, when response categories are congruent, cortical inhibition is not required, resulting in fast IAT response times. However, when the items are incongruent, the preconscious response behaviour must be over-ridden by neocortical systems, resulting in a delayed response which generates the IAT effect. This conclusion is supported by two key sources of evidence: 1) people with prefrontal damage show an inability to suppress preconscious responses (Diamond, 1998) and, 2) fMRI studies (e.g. Aron et al., 2004) have found that activation of the PFC and ACC occurs most strongly during unmatched stimulus IAT responses\textsuperscript{71}. Thus, it can be concluded that the IAT measures stimulus evaluation, or stimulus appraisal, and the IAT effect occurs as a consequence of the interaction between preconscious and cognitive appraisal mechanisms.

The IAT and stress

\textsuperscript{70} This is posited to occur via a process of stimulus evaluation (valence, relevance for attention etc) and with respect to the target category (Cunningham et al., 2004).

\textsuperscript{71} It should be noted that this observation has not been reliably replicated (e.g. Mitchell et al., 2008) and that other explanations exist, other than the cortical inhibition theory put forward here, to explain the IAT effect. For example, the IAT effect has been posited to occur as a consequence of preconscious systems needing to access semantic memory in order to categorise the stimuli accurately (Mason & Macrae, 2004). Although this is a fundamentally different proposal to Chee’s, it still explains the IAT effect in terms of the interaction between preconscious appraisal systems and cortical regulatory processes, which is the position I adopt here. Thus, to some extent, regardless of the underlying mechanism, a general conclusion may be that the IAT assesses stimulus appraisal processes.
To date, comparatively few studies have investigated the relationship between preconscious stressor appraisal and resultant stress. Those that exist tend to make use of attentional orientation tasks, such as Posner’s test (Pilgrim et al., 2010) or visual probe techniques (Fox et al., 2010), which have proven pivotal in establishing the role of preconscious appraisal processes in directing conscious attention towards stressors. For example, Ellenbogen and colleagues (Ellenbogen et al., 2010) showed that participants who had been pre-exposed to a Trier Social Stress Test (TSST) exhibited a stronger attentional bias towards later stressor stimuli than controls. Furthermore, they found that the extent of the attentional bias was strongly predictive of the magnitude of the participants’ cortisol response - evidence that robustly supports the notion that stress is attributable to the moderating effect of preconscious appraisal processes on cognitive appraisal via changes in attention. This conclusion is further supported by Jansson and Najström (2009) who observed, not only is the stressor attentional bias phenomenon replicable in other stressor-exposure procedures (in this case a Stroop-type task), but that the bias is also predictive of the participants’ autonomic reactivity to the stressor. In other words, stress (and particularly stress viewed in response-based terms, see McEwen, 2000; 2007) is strongly affected by changes in our systems of stressor appraisal.

---

72 For a review see Chapter 2.
73 Jansson and Najström (2009) used their research to develop a model of clinical anxiety based on maladaptive positive feedback within systems of attentional bias. This model proposes, in effect, that cognitive appraisal processes become sensitized to stressors by chronic alternations in preconscious appraisal processes. See also Philips (Phillips et al., 2010), who expound a similar theory for the generation of depression via changes in implicit appraisal processes.
However, whilst this research constitutes a significant step forward in our understanding of the role of stressor appraisal in eliciting stress, the major limitation of these studies (Fox et al., 2010; Pilgrim et al., 2010; Ellenbogen et al., 2010; Jansson & Najström, 2009) is that the underlying nature of the preconscious appraisal itself was not considered in their methodologies – their work only examined the effect of stressor exposure on the implicit regulation of conscious attention. In principle this limitation could be at least partly overcome using different implicit appraisal methodologies, such as the IAT. However, despite this advantage, very little work to date has employed the IAT in stress research and I have been unable to identify any studies that have used the IAT to measure stressor appraisal directly. This is surprising because research by Sato and Kawahara (2012)\textsuperscript{74} and Egloff and Schmukle (2002)\textsuperscript{75} clearly demonstrate that the IAT is sensitive enough to measure differences in implicit stressor appraisals under varying conditions. Therefore, in order to readdress this gap in the research, I intend to use the IAT to measure implicit stressor appraisal. However, in order to do so, I need first to develop and then validate an IAT instrument specific for the assessment of the appraisal of academic stressors, which will be accessible and appropriate for the adolescent participants taking part in this research. This goal constitutes the rationale for this Chapter.

**Developing the Stressor Appraisal IAT (SA-IAT)**

---

\textsuperscript{74} Sato and Kawahara (2012) used the IAT to test for the effect of stress on the categorisation of 'self'/other' stimulus items and, thereby, confirm the existence of acute stress in the participants (although see Schmukle & Egloff, 2004).

\textsuperscript{75} Egloff and Schmukle (2002) used the IAT to establish the existence of stress in order to validate a self-report questionnaire they were developing.
During IATs participants undergo a simple behavioural response-based learning program in which they are asked to classify stimulus words into single and then paired combinations of super-ordinate categories. Although a variety of IAT designs now exist (such as the Brief Implicit Association Test – see Sriram & Greenwald, 2009) the majority of IAT procedures adopt the traditional five-step procedure designed by Anthony Greenwald (Greenwald et al., 1998), which can be summarised as follows.

Step 1: This is an initial target-discrimination step in which the participants allocate stimulus words to one of two super-ordinate target categories: in this case 'exams' (a blanket term for academic stressors) and an appropriate antonymous category (here 'holidays' was chosen). This step is designed to familiarise participants with stimulus items, the key-pressing procedure and the super-ordinate concept categories.

Step 2: This step introduces the stimulus attributes and reinforces the participants' familiarity with the IAT procedure. In a similar two-category discrimination task, participants allocate stimulus words into one of two antonymous super-ordinate attribute categories (in this case 'stressed' and 'relaxed').

Step 3: In this step both target and attribute categories are superimposed (i.e. 'exams'+'stressful' and 'holidays'+'relaxing'). Stimulus words for target and attribute discriminations appear on alternating trials. The speed at which participants allocate individual stimulus words into the appropriate paired category is recorded.
Step 4: This step is a repeat of Step 1 but with the key press response assignments reversed. This step prepares participants for Step 5 and also controls for unintentional directional bias or preference for response assignments.

Step 5: This step is a repeat of Step 3 but with the target and attribute categories paired the other way around (i.e. ‘exams’+‘relaxing’ and ‘holidays’+‘stressful’). As with Step 3, response speeds are timed and will later be compared with those of Step 3. An IAT effect is generated if the stimulus categorisation timings differ between Steps 3 and 5. If the target categories are similarly associated with the attribute dimension, participant tend to categorise the same list of stimuli (it is identical in Steps 3 and 5) with equal speeds in both Step 3 and Step 5. However, if a participant responds significantly faster on one of the two steps - i.e. the participant finds this step considerably easier than the other - the implication is that an implicit association exists between the target and attribute categories (Greenwald et al., 1998).

As the majority of published studies adopt the five-step procedure outlined above, I chose to base the Stressor Appraisal IAT (SA-IAT) on Greenwald’s (Greenwald et al., 1998) original design, but using his improved D scoring algorithm (Greenwald et al., 2003). A number of software programmes are available which allow for the construction and scoring of IATs using these parameters (e.g. the Free IAT design software developed by Meade [see Meade, 2012], or the IAT Design package [see Jander & Jander, 2010]). However, the most widely used is the Inquisit IAT software developed by Millisecond (see Millisecond Website), which has been utilised by over a thousand research groups,
including Greenwald's (e.g. Dasgupta et al., 2000). The Inquisit software allows not only for the SA-IAT instrument to be completed on-line (via a web hosting or, in this case, using a school VLE platform), but also for multiple participants to complete the SA-IAT concurrently, saving considerable time. A further benefit is that the Inquisit software includes a high precision key-press timing system, so that response times to Steps 3 and 5 are automatically recorded (and any difference calculated), thus removing with the need to record IAT sessions digitally\textsuperscript{76} and calculate key-press timings retrospectively by playing back recorded sessions at half-speed. However, in developing the SA-IAT using the stipulations of the Inquisit software, two key decisions had to be made in order to produce the SA-IAT instrument, namely; 1) the choice of appropriate antonymous terminology for the target and attribute categories and, 2) choice of the stimulus words.

1. In developing the High School Stressors Scale, Burnett and Fanshawe (1997) identified ‘examinations’ as the most significant stressor adolescents experience at school\textsuperscript{77}. This finding replicated Kohn and Frazer's (1986) earlier conclusion that examinations were the most significant contributor to school-based stress. Thus, I chose ‘exams’ as the blanket term for the academic stressor target category. Although this choice was obvious, it presented an unforeseen problem in the choice of antonymous term for the partner target category: i.e. ‘exams’ has no immediately obvious antonym. To overcome this problem I conducted a brief survey of Sixth Form students to establish an appropriate word

\textsuperscript{76} And the ethical considerations associated with such practices (see Mertens, 1998).

\textsuperscript{77} Their study demonstrated that exams correlate with 69% of self-reported stress - making it by far the most significant stressor.
that participants would readily associate as semantically opposite to 'exams'. In the survey, 33 female participants (all between 16 and 18 years of age) were each given 20 seconds to list antonyms for the word 'exams'. Their responses were collated and ranked according to their frequency to determine popularity. The data collected demonstrated overwhelmingly that participants associate the word 'holidays' as antonymous to 'exams' and, therefore, 'holidays' was selected as the second target category to partner 'exams'.

Choice of the attribute terms proved more straightforward. I chose 'stressful' as the most appropriate term to describe the ability of a stressor to elicit stress and, as the Shorter Oxford English Dictionary lists 'relaxing' as the primary antonym for 'stressful', the attribute terms 'stressful' and 'relaxing' were selected for the SA-IAT.

2. Specific words have strong affective and emotional connotations (Ainsfeld & Lambert, 1966; Teasdale & Russell, 1983) and, in order to avoid obfuscating the implicit target-concept association, it is critical to select stimulus items for the SA-IAT that possess equivalent semantic 'value'. Additionally, stimulus items must be readily comprehended by the participants (i.e. be words within their common vocabulary) in order for implicit associations to be revealed by the SA-IAT. Therefore, in order to generate a list of appropriate stimulus items, a second pilot study was conducted using the same group of 33 female adolescent participants, which followed a broadly similar method to that employed by Belezza, Greenwald and Banaji (1986) in that it required participants to complete a Semantic Categorisation Questionnaire instrument. By completing this instrument, participants confirmed not only whether potential stimulus
items were in their common usage, but also helped to generate approximate semantic normative values for the stimulus items, which then enabled me to select items for the SA-IAT that were both semantically equivalent and in the participants’ common parlance.

The Semantic Categorisation Questionnaire (SCQ) instrument employed in the pilot was constructed from words selected from lists included by Russell (1980) in his emotion circumplex model. Twenty items synonymous for 'stressful' were taken from the extreme perimeter of the 'obstructive-negative' quadrant of his affective semantic wheel and a further twenty items synonymous for 'relaxing' were selected from the 'conducive-positive' quadrant. Russell's (1980) emotion circumplex model attempts to map affective vocabulary in two-dimensional space according to valence and activity/arousal (Russell, 1983). Thus, in selecting words from the similar areas of the wheel, the affective valence and arousal of items is effectively controlled, ensuring similar semantic value. Having selected words for the SCQ, participants completed the instrument to indicate (using a 5-point Likert-type scale) the degree to which they found the stimulus items 'relaxing' or 'stressful'. This served to confirm the semantic valuation and also to ascertain the familiarity of the word within the participants' vocabulary (participants were instructed not to score unfamiliar words). From the SCQ scores, the top and bottom ten items were chosen for inclusion in the SA-IAT.

Therefore, having found solutions to the questions outlined above, the design of the SA-IAT was completed using the Inquisit software and is represented in Figure A2 below.
Establishing the validity of the SA-IAT

After the publication of the original IAT in 1998, a large body of research accumulated to support the validity of the instrument, which led to the IAT receiving rapid and extensive support (Devine, 2001). This espousal owed, at least in part, to studies demonstrating a high degree of reliability in the IAT instrument, both internally (with alpha equal to approximately 0.8 - see Greenwald & Nosek, 2001) and externally via test-retest procedures (Greenwald et al., 2002). Similarly, positive measures of temporal concordance have been recorded (e.g. Cunningham et al., 2001 who recorded a stability of .46 over two weeks) supporting earlier research upholding the stability of IAT instruments (Dasgupta & Greenwald, 2000; Bosson et al., 2000; Egloff et al., 2005; although see Gschwender et al. [2008] who found that IATs could be affected by context).

---

**Figure A2**: Schematic illustration of the Stressor Appraisal Implicit Association Test (adapted from a similar diagram in Greenwald et al., 1998)

<table>
<thead>
<tr>
<th>Step</th>
<th>Task Description</th>
<th>Task Instruction</th>
<th>Sample Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initial Target-Concept Discrimination</td>
<td>• EXAMS • HOLIDAYS •</td>
<td>VACATION • TEST • ASSESSMENT BREAK • MOCK REST • GOING ABROAD • QUIZ •</td>
</tr>
<tr>
<td>2</td>
<td>Associated Attribute Discrimination</td>
<td>• Stressful Relaxing •</td>
<td>• Frequent Jitter Tranquil • Nervous Calm • Strained Peaceful •</td>
</tr>
<tr>
<td>3</td>
<td>Initial Combined Task</td>
<td>• EXAMS • Stressful HOLIDAYS • Relaxing •</td>
<td>• Paper Tense WEEKEND AWAY • Chilled • TIME-OFF • Worried • FINALS Soothed •</td>
</tr>
<tr>
<td>4</td>
<td>Reversed Target-Concept Discrimination</td>
<td>EXAMS • HOLIDAYS •</td>
<td>• TRIP SOJOURN CHECK • LEAVE INVESTIGATION QUALIFICATION • FREE-TIME MODULES •</td>
</tr>
<tr>
<td>5</td>
<td>Reversed Combined Task</td>
<td>EXAMS • HOLIDAYS • Relaxing •</td>
<td>• Overwrought RELAXATION PERIOD • Placid • APPRAISAL • Hassled STAYING HOME • Cool • TASK •</td>
</tr>
</tbody>
</table>
Support was further bolstered when IAT effects were shown to be independent of handedness bias (Greenwald & Nosek, 2001), assignment of categories to specific sides (Greenwald et al., 1998) intertrial-interval duration (Greenwald et al., 1998), order of combined tasks (Nosek et al., 2005) variation in the number of stimulus items, familiarity of items\(^{78}\), variability in the duration of the response-stimulus interval, or the order of mixed categorisation tasks (reviewed in Brunel et al., 2004).

However, despite this early positive reception, demonstrations of reliability verify only content validity, which accounts for only a slim portion of construct validity (Messick, 1995) - a point that did not go unnoticed as reviews of the IAT (e.g. Cunningham et al., 2001; Kihlstrom, 2004) published an urgent call for further research to address key questions over the use of the IAT as a valid measure of implicit cognition. This call for inquiry prompted a flurry of research, which Greenwald and his colleagues (Nosek et al., 2007; Lane et al., 2007) then reviewed extensively. Their meta-analyses concluded that the validity of the IAT had largely been established insomuch that it could be demonstrated in all three of the subcategories of validity identified by Messick (1989) - i.e. Content validity (Reliability, or Internal validity), Criterion-related validity (also known as Concurrent or Predictive validity - Brown, 1996) and Construct validity - as well as a fourth category of ecological, or Consequential Validity, which is the aspect of construct validity that “appraises the value implications of score interpretation as a basis for action” (Messick, 1995). I briefly outline this evidence below.

---

\(^{78}\) Except in circumstances where familiarity of stimulus items is very low (Brendl et al., 2001).
1. **Construct Validity.** Cooper and Schindler (2001) advocate the use of convergent and discriminant measures in establishing the validity of a construct, such as implicit appraisal theory. Applied to this field, construct validity is most commonly demonstrated through the convergence of conclusions based on data taken from separate implicit measures run concurrently across the same group of participants. For example, in a study on self-esteem Bosson, Swann and Pennebaker (2001) compared seven different implicit measures (including the IAT) and found that these instruments produced broadly supportive results. Other studies have compared IAT instruments with priming procedures (Sherman et al., 2003; Marsh et al., 2001), where convergence was also established. Criticism of these studies highlights the finding that these validating supportive relations are often weak, which might seem to undermine the principle of convergence. However, in a review of such studies Greenwald explains that:

"Implicit measures often demonstrate relatively weak reliability compared to other forms of psychological measurement. Reliability of measures set upper limits on their possible correlations with other measures. For example, the maximum, meaningful correlation that can be observed between a measure with reliability of .10 and a measure with perfect reliability (1.0) is .32, which is estimated by calculating the product of the square roots of the two reliability coefficients" (Nosek et al., 2007, pg277).

Therefore, given that the IAT and other implicit measures can have relatively low reliability (compared to other quantitative measures), weak correlations between implicit measures actually corresponds to a strongly convergent measure of validity. This becomes more obvious when the low initial reliability of some implicit measures is compensated for...
statistically, which Cunningham and colleagues (Cunningham et al., 2001) and Nosek and Banaji (2001) have made the focus of independent research projects. Their work demonstrated significantly higher correlations between implicit measures (up to 0.77) which, taken together, constitutes strong evidence of construct validity via convergence, thus adding significantly to the validity of the IAT as a measure of implicit appraisal.

Similar to convergence studies, triangulation methodologies employ separate qualitative and quantitative data collection processes “to obtain different but complementary data on the same topic” (Morse, 1991, cited in Creswell & Plano Clark, 2007, pg87; see also Creswell, 1999). This approach has been adopted in IAT research, where studies have attempted to establish validity by analysing the correlation between IAT scores and parallel explicit self-report measures. This approach initially met with mixed success: whilst some studies reveal high correlations (Nosek et al., 2002a), others show no significant correlation at all (Nosek & Banaji, 2002). Such contradictory findings prompted wide-scale meta-analyses (Nosek, 2005; Poehlman et al., 2004; Hofmann et al., 2005) which determined that, whilst explicit and implicit measures do show a high degree of inter-relation, they should be considered as distinct constructs (Nosek, 2005; Nosek & Smyth, 2007). This conclusion is not surprising because neuroimaging studies (Cunningham et al., 2003; 2004; Phelps et al., 2000), psychometric evidence (Greenwald & Farnham, 2000; Nosek & Smyth, 2005) and dual process theories of emotion79 all delineate explicit and implicit appraisal processes as being neurologically and functionally separate. Apropos, some researchers have concluded

---

79 See Chapter 3 (see also Wilson et al., 2000 and the opening section of this Chapter).
that studies demonstrating an absence of (or low) correlation between implicit and explicit techniques (Nosek & Smyth, 2007; Nosek & Hanson, 2008) constitute evidence of construct validity via discriminant technique (Cooper & Schindler, 2001) - i.e. when similar instruments fail to converge, rightly, because they measure separate constructs. However, although this argument initially seems compelling, on inspection it proves paradoxical as one cannot take the presence of correlations between implicit and explicit methods as evidence for construct validity (Poehlman et al., 2004; Hofmann et al., 2005) whilst also arguing that the absence of such convergence also constitutes evidence of construct validity by discriminant technique (Nosek, 2005; Nosek & Smyth, 2007). This impasse has, in part, been created through the employment of a diverse array of implicit and explicit instruments across the literature, each of which allows for a different blend of interaction between explicit and implicit appraisal processes. This is further confounded by the interdependence of the appraisal processes themselves\(^80\). Therefore, whilst data from parallel explicit and implicit measures does, for the most part, support the construct validity of the IAT (Nosek, 2005; Poehlman et al., 2004; Hofmann et al., 2005), further work is certainly required to establish this beyond reasonable doubt.

2. Predictive Validity. Studies of convergence analysis have brought to light another possible avenue for validation of the IAT, namely the potential for implicit measures to predict behaviour (Maison et al., 2004). A number of early studies had shown positive correlations between IAT measures and the prediction of maladaptive behaviours, including anxiety (Asendorpf et al., 2002) and chronic alcohol dependence (Wiers et al.,

\(^{80}\) See Chapter 3.
2002). However, Nosek’s (Nosek et al., 2002b) observation that implicit negativity towards mathematics was heavily correlated with performance in maths SAT examinations suggested that the IAT might also prove predictive of mainstream behaviours. In a meta-analysis of 122 studies in this field, Greenwald (Greenwald et al., 2009) concluded exactly this, particularly for socially sensitive topics, where the IAT significantly outperformed alternative explicit instruments. Although he advocates a combination of explicit and IAT measures as the best predictor of behaviour, Greenwald’s (Greenwald et al., 2009) finding that IAT scores are correlated with behavioural probabilities does achieve the effect of methodological triangulation and, therefore, adds credence to the application of the IAT as a valid measure of implicit stressor appraisal.

3. Internal Validity. Response-latency studies are not renowned for having high internal reliability (Lane et al., 2007) as error variance is too easily introduced across items (e.g. by participant distraction). Equally, the internal consistency of the IAT is frequently measured as being lower that that of self-report instruments (Buchner & Wippich, 2000). Taken together, this evidence could lead to the inference of low reliability within the IAT. However, that conclusion is not supported by the wider literature as meta-analytic studies (e.g. Hofman et al., 2005) place average alpha scores at .79 for the IAT, indicating a relatively high level of consistency across the instrument. This is in keeping with the observation that, of the available measures of implicit appraisal, the IAT is consistently recorded as having the best internal reliability (Cunningham et al., 2001). Test-retest procedures also support this view as, in a study of implicit appraisal of self-esteem, Bosson and colleagues (Bosson et al., 2004) found that the reliability of implicit instruments was
significantly higher for the IAT across successive tests than that of other implicit measures. Other test-retest literature also lends support to the internal validity of the IAT as a number of studies have demonstrated strong reliability correlations between test and retest (Dasgupta et al., 2000; Egloff et al., 2005; Steffens & Buchner, 2003)\(^1\), lending support to the conclusion that the IAT shows a high degree of reliability.

Despite this evidence, some researchers have identified specific threats to internal validity within IAT measures. Of these, familiarity of items has attracted the most research. Citing Zajonc's (1968; 2001) observation that familiar stimuli tend to be favoured in preference to unfamiliar stimuli, Rothermund and Wentura (2001) demonstrated that IAT-like effects could be produced in a 'go/no-go' task simply by manipulating the salience of the target-descriptor category. This led them to propose that the IAT effect occurs through asymmetry of stimulus valence (Rothermund & Wentura, 2004), rather than through implicit stimulus evaluation as Greenwald advocates (Greenwald et al.; 1998; 2003; 2009) - a clear challenge to the internal validity of the IAT.

Rothermund and Wentura's argument hinges on the assertion that participants tend to focus predominantly on one of the two IAT categories, effectively remodelling the IAT task as a stimulus-detection activity in which salience asymmetry wholly governs response time (either directly through familiarity via pre-exposure, or indirectly through natural sources of asymmetry [e.g. linguistic markedness (Lyons, 1977), or

\[^1\] However, see also Dasgupta and Asgari (2004), who found very low test-retest reliability correlations.
the popout effect (Strayer & Johnson, 2000). This position is supported by some experimental evidence. For example, Brendl and colleagues (Brendl et al., 2001) demonstrated that positive and negative IAT effects could be reversed when 'nonwords' were included as stimulus items, implying that the lack of familiarity (with 'nonwords') drove the IAT effect they had observed. Their conclusion that item familiarity confounds the IAT effect is also supported by Rothermund and Wentura (2001), whose 'association-free' IAT designs were found to elicit effects similar to those measured by 'normal' IAT procedures. Similarly, inverting salience asymmetry of stimulus items has been shown clearly to reverse the IAT effect (Rothermund & Wentura, 2004), whereas manipulation of association did not alter the IAT effect in the same experiments if salience asymmetries were held constant - outcomes that not only support Rothermund and Wentura's (2004) theory of salience asymmetry, but also appear to undermine considerably the internal validity of the IAT instrument.

However, counterbalancing this data is large body of evidence suggesting that stimulus familiarity has little or no effect on the IAT. For example, in a key study, Dasgupta and colleagues (Dasgupta et al., 2000) controlled stimulus familiarity by use of a statistical regression and still observed a marked IAT effect across participants - a result that directly contradicts Rothermund and Wentura's salience asymmetry account. Similarly, in unconscious stereotyping experiments, the use of unfamiliar stimulus items led to the robust preservation of IAT responses (Ottaway et al., 2001, see also Rudman et al., 1999), which is also at odds with the notion of familiarity-based stimulus processing. Finally, it has been shown that participants' familiarity with words and symbols related to
mathematics does not contribute to an 'automatic liking' for that discipline as measured by the IAT (Nosek et al., 2002b). Therefore, the overall conclusion must be that salience symmetry, by itself, does not generate the IAT effect. However, given the number and variety of studies that do support a role for stimulus valence in IAT processing, I accept Kinoshita and Peek-O’leary’s (2005, pg442) reservation that IAT effects can also be “interpreted in terms of salience asymmetry, based on dimensions like familiarity and linguistic markedness.” Therefore, whilst the IAT does display significant internal validity (Hofman et al., 2005), I acknowledge that a potential limitation in its use is the confounding effect of stimulus asymmetry. As yet, a solution to this issue has not been proffered, so I await future research clarifying this area with enthusiasm and interest.

Aside from stimulus familiarity, four other factors have been shown to have a limiting effect on the internal validity of the IAT: namely, a) participants’ familiarity with the IAT itself, b) the fakeability of the IAT effect, c) the age of participants and, d) cognitive fluency of participants.

a) Pre-exposure to implicit instruments has been shown to have a small, but significant, effect on the IAT (Greenwald et al., 2003). Although an interpretation of this result has yet to be offered, it nevertheless constitutes an additional threat to internal validity. However, given the age and demographic of the participants in this study, the chance of any of them having previously completed an IAT is miniscule and, therefore, presents little threat to the internal validity of the SA-IAT as employed here.
b) Despite Kim's (2003) evidence that the IAT effect is not deliberately fakeable, the consensus of studies tend to draw the opposite conclusion (e.g. Fiedler & Bluemke, 2005; Gregg et al., 2006; De Houwer et al., 2007), leading to the suggestion that IAT tests are easily biased by demand compliance (De Houwer, 2006) and could, therefore, be unreliable. However, whilst I cannot rule out this possibility, most evidence of fakeability derives from studies in which novel words (Mitchell et al., 2003) or 'nonwords' (Gregg et al., 2006) are used and, as De Houwer and colleagues report (De Houwer et al., 2007), this result is difficult to replicate in IATs with well-known attitude objects such as those utilised in the SA-IAT. Additionally, as the ability to fake an effect increases with experience of IAT (Steffens, 2004; Cvencek et al., 2010), demand compliance is doubly unlikely to have a significant biasing effect on response timing, in either this research, or that of the main study. Therefore, I posit that the influence of fakeability on the internal reliability of the SA-IAT is also minimal.

c) Response latency measures are well known to suffer from age-related biases (Faust et al., 1999; Ratcliff et al., 2000) and the IAT is no exception (Greenwald & Nosek, 2001; Hummert et al., 2002). However, not only has Greenwald significantly adapted the IAT scoring algorithm to compensate for this effect (Greenwald et al., 2003), but the participants contributing to this research and the main study have all been selected from the same school year (the Sixth Form), ensuring that their age was controlled and,
therefore, that any age-related threats to internal validity were eliminated.

d) In keeping with the previous point, response-based measures have been shown to be biased both by intelligence (Jensen, 1993; Salthouse, 1996) and the cognitive fluency of participants to cope with switching between intellectual processing tasks (Salthouse et al., 1998). These influences have been shown to affect IAT measures as McFarland and Crouch (2002) were able to demonstrate that independent IAT scores (measuring very different association processes) could be made to correlate closely with each other: a result that led McFarland and Crouch (2002) to conclude that underlying cognitive ability is responsible for producing the IAT effect. However, whilst Greenwald acknowledges this influence (Cai et al., 2004), he also clearly demonstrates that cognitive fluency only confounds millisecond-measures of latency, such as those utilised in McFarland and Crouch’s study. Greenwald has since presented a new scoring algorithm for the IAT (Greenwald et al., 2003) which rescales individual IAT effects according to within-participant latency variability, which carefully controls for individual variations in cognitive fluency as well as intelligence. Greenwald (Cai et al., 2004) reports that, when this algorithm is employed in IATs, the biasing influence detected by McFarland and Crouch (2002) is reduced to insignificant levels\textsuperscript{82}. This observation has been replicated by others (Mierke & Klauer, 2003), supporting the view

\textsuperscript{82} Although, note, that the influence of intelligence and cognitive fluency is not eliminated entirely by Greenwald’s (Greenwald et al., 2003) new algorithm.
that neither cognitive fluency, nor intelligence significantly confounds IAT effects and thus that the IAT demonstrates robust internal validity\textsuperscript{83}. Therefore, with the caveat of stimulus asymmetry (discussed above), the evidence presented here strongly supports the conclusion that the SA-IAT, as employed in this research, constitutes a \textit{reliable} implicit measure.

4. \textbf{Consequential Validity}. Blanton and Jaccard (2006) have criticised the IAT on the grounds of low \textit{metric meaningfulness}, arguing that it fails to reveal the underlying psychological dimension it was designed to measure (implicit associations between stimuli). Because of this, Blanton and Jaccard (2006) assert that the IAT constitutes an \textit{arbitrary} analytical instrument and thus, in keeping with Shepard’s (1997) and Popham’s (1997) interpretation of consequential validity,\textsuperscript{84} they argue that the IAT must be considered as ecologically invalid because it leads researchers towards unintended and inappropriate inferences. Blanton and Jaccard (2006) cite the absence of normalisation (i.e. measurement relative to normative values) and the lack of a fixed zero-point as specific examples of the lack of meaningfulness in the IAT. However, in regard to their first point, Nosek and Sriram (2007) observe that, in order to conform to a normalised scale, relative preferences between target concepts and relative attributes must interact \textit{additively} in order to form a framework for comparison between concept-attribute pairs. This is not the case “as relative preferences are not decomposable into component attitudes at all” (Nosek & Sriram, 2007, pg394); rather they are interdependent

\textsuperscript{83} See also Stüttgen et al., 2011, who argue for a different IAT scoring process to address a flaw in Greenwald’s (Greenwald et al., 2003) scoring correction.
\textsuperscript{84} See also Messick (1995).
(Hsee et al., 1999) and demonstrate multiple levels of additivity (Anderson, 1977). Although measures of single association are now available (e.g. De Houwer’s ‘EAST’ (the Extrinsic Affective Simon Task - see De Houwer, 2003), they have yet to be validated and, as the IAT is not an appropriate measure for single association (Nosek et al., 2005), there is limited scope to accommodate for Blanton and Jaccard’s (2006) call for the IAT to be normalised. This remains a limitation in the use of the IAT in research and, potentially, a threat to consequential validity if utilised by researchers with a lack of awareness of the limitations of the IAT.

By contrast, Blanton and Jaccard’s (2006) second criticism - that the IAT contains an implicit assumption that the IAT zero-point maps onto neutral stimulus association, and that this has yet to be validated - has been robustly dismissed. In a swift response to Blanton and Jaccard’s 2006 paper, Greenwald (Greenwald et al., 2006) provided succinct evidence of zero-point congruence by comparing explicit and implicit preferences for presidential candidates. Their work clearly demonstrated IAT scores centring about a zero-point, which correlated very closely with the neutral position of the self-reports. Therefore, whilst I accept Blanton and Jaccard’s (2006) criticism that the IAT has not been normalised (and thus that measures of relative implicit association recorded in this research must be considered as exactly that - relative measures rather than absolute measures), I reject their assertion that the IAT lacks meaningfulness. As the IAT has been shown to demonstrate predictive validity in 86 separate and independent studies (Poehlman et al., 2004), I posit that the meaningfulness of the IAT has indeed been established (albeit from a relativistic perspective) and that,
in conjunction with the evidence presented above, that the SA-IAT constitutes a thoroughly and demonstrably valid measure of implicit stimulus association (Nosek et al., 2007; Lane et al., 2007).

Using the Stress Appraisal Measure to validate the SA-IAT

As I have demonstrated above, IAT instruments comprise valid measures of implicit association. However, despite their inventor's repeated affirmation that IATs measure differences in preconscious appraisal processes (e.g. Greenwald et al., 1998; 2009; Greenwald & Nosek, 2001), other contending theories have also been advanced. For example, IAT effects have been ascribed to environmental associations (Karpinski & Hilton, 2001), response criterion shifts (Brendl et al., 2001), task-set switching (Mierke & Klauer, 2001; 2003), response conflicts (De Houwer, 2001) and stimulus valence asymmetry (Rothermund & Wentura, 2001; 2004), to mention but a few. This diversity of discourse undermines the validity of the SA-IAT as a measure of implicit stressor appraisal. Added to this - as mentioned previously - I have been unable to identify any studies within the literature where IATs have been utilised specifically as measures of implicit stressor appraisal, which further limits the construct validity of the SA-IAT. Therefore, in order to justify the use of the SA-IAT in the manner I propose, I chose to conduct a small research project to validate the SA-IAT as a measure of implicit stressor appraisal.

The study was designed to follow a method similar to that employed by Nosek and his colleagues (Nosek et al., 2002a) in that it utilised concurrent implicit and explicit instruments in an attempt to establish
concordance between (and thus the validity of) similar measures. The implicit measure chosen for this study was, of course, the SA-IAT, whilst the second instrument selected was the Stress Appraisal Measure (SAM - Peacock & Wong, 1990, see Appendices C & D) - a well-used explicit measure of stressor appraisal. Briefly, the SAM is a 28-item, 5-point Likert-type instrument\textsuperscript{85}, initially developed by Peacock and Wong (1990) to assess the appraisal of an anticipated academic stressor. The SAM uses a multi-dimensional, seven sub-scale scoring system to model stressor appraisals in terms of Lazarus's cognitive-relational theory (Lazarus, 1991; 1999; Lazarus & Folkman, 1984), either as 'threat'/'challenge' primary appraisals, or a variety of secondary appraisal processes. Although the SAM is an explicit measure (and thus accesses different, though interlinked, neurological appraisal processes to the SA-IAT [Cunningham et al., 2003; 2004]\textsuperscript{86}), it is an appropriate instrument to use in this study because the SAM is “one of the few instruments... specifically designed to assess dimensions of primary and secondary appraisal separately [my emphasis] and which explicitly attempts to distinguish coping processes from appraisal processes” (Cohen et al., 1997, pg32). As Lazarus’s (1991) ‘primary appraisals’ have been defined either as being similar to implicit appraisal processes (Lazarus, 1999; Ochsner & Barrett, 2001), or identical to them (Marsella & Gratch, 2009), by distinguishing between primary and secondary appraisal processes, the SAM enables us, to a point, to affect a measure of control on the influence of explicit processes during the completion of the SAM instrument. This, therefore, allows for legitimate, quantitative comparison between the implicit SA-IAT data and the quasi-implicit

\textsuperscript{85} For a critique of Likert-type measures please see Chapter 4.

\textsuperscript{86} See also the section of this Chapter on Construct Validity.
primary appraisal scores generated from the ‘threat’ and ‘challenge’ sub-scales in the SAM\textsuperscript{87}.

Furthermore, as well as the ability to limit the effect of explicit appraisals, a further justification for the use of the SAM to validate the SA-IAT is that the SAM instrument robustly demonstrates high internal validity (α ranges from 0.74 to 0.90 - see Roesch & Rowley, 2005) and, because of its multidimensional structure, it has been employed in a wide variety of contexts (Anshel et al., 1997; Peacock & Wong, 1990; Roesch & Rowley, 2005; Senol-Durak & Durak, 2012), which has the further advantage in that a number of robust assessments of validity have been completed. These include test-retest research (Peacock & Wong, 1990; Durak, 2007) and convergent and discriminant procedures (Rowley et al., 2005; Roesch & Rowley, 2005). Taken together, these studies confirm Peacock and Wong’s (1990) initial assertion that the SAM is a valid and reliable measure of academic stressor appraisal and, therefore, an appropriate instrument to use to validate the SA-IAT as a measure of implicit stressor appraisal.

**The IFS Study**

Thirty-seven participants volunteered to take part in the study. They were split into two groups (a group of 18 and a group of 19); one group completed the SA-IAT followed by the SAM and the other half completed the instruments in the reverse order. However, as three

\textsuperscript{87} Although Peacock and Wong’s (1990) original SAM instrument splits primary appraisals into ‘threat’, ‘challenge’ and ‘centrality’ sub-scales, the ‘centrality’ scale has been found to work poorly with adolescents (Rowley et al., 2005) and will be not be scored in this study. Thus only data from the ‘threat’ and ‘challenge’ sub-scales will be used here.
participants left sections of the SAM blank, and two students failed to return their consent forms with parental counter-signatures, only 32 sets of data went forward for analysis.

Based on Creswell's (2003) guidance, before taking part in the study, all participants were given a full briefing, which explained their right to withdraw from the study at any point, their right to withdraw their data from the research if they so desired, and their right to ask questions and to obtain a copy of the results and an outline of the benefit of the study. The participants were also asked to sign a consent form indicating that they understood the procedure and that they had given their informed consent to take part. As the majority of students were under the age of 18, all of the participants were asked to give their briefing and consent forms to their parents for counter-signing to indicate parental consent for participation. As stated previously, two students failed to return counter-signed forms and were, therefore, not allowed to participate. Additionally, in order to maintain participant confidentiality (but allowing completed SAM & SA-IAT instruments to be paired to their respective respondents), each participant was instructed to follow a procedure that assigned them a unique, anonymous pseudonym - the first two letters of their mother's maiden name and the last two digits of their phone number. This is in keeping with strategies recommended by Frankfort-Nachmias and Nachmias (1992, cited in Cohen et al., 2008). Data from the study will be kept for 5 years (as recommended by Sieber, 1998, cited in Creswell, 2003).
### Data.

<table>
<thead>
<tr>
<th>Participant</th>
<th>SAM Scores</th>
<th>Threat Qs</th>
<th>Challenge Qs</th>
<th>SA-IAT D Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5 11 20 28</td>
<td>Mean 7 8 10 19</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>4 3 4 3</td>
<td>3.50 2 2 3 4</td>
<td>2.75 1.37</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>4 3 3 3</td>
<td>3.25 2 2 2 3</td>
<td>2.25 1.15</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>4 4 3 3</td>
<td>3.50 3 2 2 3</td>
<td>2.50 1.21</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>3 2 4 4</td>
<td>3.25 3 1 3 2</td>
<td>2.25 0.88</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>4 2 3 3</td>
<td>3.00 2 2 2 2</td>
<td>2.00 1.11</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>3 5 4 3</td>
<td>3.75 2 2 2 3</td>
<td>2.25 0.92</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>2 3 3 2</td>
<td>2.50 3 2 2 2</td>
<td>2.25 0.67</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>4 3 4 3</td>
<td>3.50 2 2 4 3</td>
<td>2.75 1.33</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>4 4 5 3</td>
<td>4.00 1 2 2 3</td>
<td>2.00 1.01</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>3 3 3 3</td>
<td>3.00 3 3 2 2</td>
<td>2.50 0.92</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>3 4 3 3</td>
<td>3.25 2 1 2 4</td>
<td>2.25 1.11</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>4 4 3 3</td>
<td>3.50 3 2 2 2</td>
<td>2.25 1.08</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>4 3 3 3</td>
<td>3.25 2 2 2 1</td>
<td>1.75 1.22</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>4 5 2 3</td>
<td>3.50 2 3 3 2</td>
<td>2.50 0.55</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>2 2 3 3</td>
<td>2.50 2 3 2 2</td>
<td>2.25 0.76</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>4 2 3 4</td>
<td>3.25 3 3 2 2</td>
<td>2.50 1.18</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>3 3 3 4</td>
<td>3.25 2 2 2 2</td>
<td>2.00 1.20</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>4 3 5 4</td>
<td>4.00 2 2 2 3</td>
<td>2.25 1.11</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>4 4 4 4</td>
<td>4.00 2 2 1 2</td>
<td>1.75 1.19</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>3 2 3 3</td>
<td>2.75 2 2 1 1</td>
<td>1.50 1.33</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>3 4 3 3</td>
<td>3.25 2 3 2 2</td>
<td>2.25 0.65</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>4 4 4 3</td>
<td>3.75 1 2 2 1</td>
<td>1.50 0.98</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>4 3 4 4</td>
<td>3.75 2 2 2 2</td>
<td>2.00 1.33</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>3 2 3 4</td>
<td>3.00 2 3 1 1</td>
<td>1.75 1.28</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>2 2 4 4</td>
<td>3.00 2 2 2 2</td>
<td>2.00 0.77</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>3 4 3 4</td>
<td>3.50 2 2 2 3</td>
<td>2.25 1.33</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>4 4 4 5</td>
<td>4.25 2 2 2 4</td>
<td>2.50 1.27</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>2 3 4 3</td>
<td>3.00 2 3 1 1</td>
<td>1.75 1.11</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>2 2 2 3</td>
<td>2.25 2 2 3 3</td>
<td>2.50 1.06</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>3 2 4 3</td>
<td>3.00 3 2 2 2</td>
<td>2.25 0.65</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>3 4 3 3</td>
<td>3.25 2 2 1 3</td>
<td>2.00 0.89</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td>4 4 3 5</td>
<td>4.00 2 2 1 2</td>
<td>1.75 1.26</td>
</tr>
</tbody>
</table>

*Table A1*: Data showing SAM question scores for ‘threat’ and ‘challenge’ sub-scales and SA-IAT D Score.
Table A2: Calculation of Cronbach's Alpha for items assessing the 'challenge' sub-scale of the SAM.

Table A3: Calculation of Cronbach's Alpha for items assessing the 'threat' sub-scale of the SAM.

Table A4: Calculation of inter-item correlation for items assessing the 'threat' sub-scale of the SAM.

Table A5: Calculation of Cronbach's Alpha for removal of items assessing the 'threat' sub-scale of the SAM.
Graphs.

**Figure A3**: Graph showing the correlation between average score on the ‘threat’ sub-scale of the SAM against participants’ SA-IAT scores.

**Figure A4**: Graph showing the correlation between average score on the ‘challenge’ sub-scale of the SAM against participants’ SA-IAT scores.
Discussion

'Challenge' Appraisals. Figure A4 demonstrates a very slight negative correlation between participants' average scores on the 'challenge' sub-scale of the SAM and their SA-IAT D scores. However, the correlation is insignificant ($r = 0.113$), even at the 10% significance level ($r_{crit} = 0.296$). Two possible conclusions can be drawn from this observation:

a) In contrast to previous studies (Peacock & Wong, 1990; Rowley et al., 2005; Roesch & Rowley, 2005), the 'challenge' sub-scale of the SAM is unreliable as a measure of 'challenge'-type primary appraisals. This conclusion is supported by Table A2, which shows the Cronbach's $\alpha$ value calculated for these items (0.005), indicating a complete absence of reliability within this sub-scale of the SAM instrument!

b) That the implicit appraisal processes measured by the SA-IAT are separate constructs from 'challenge' appraisals and thus no concordance of measures was observed. On the face of it this conclusion appears relatively straightforward as Lazarus (1984; 1991) and other early appraisal theorists (e.g. Leventhal & Scherer, 1987) appeared to define appraisals primarily as 'cognitive' events, implying that implicit processes play no (or a very limited) part in 'challenge' appraisals. However (and without wishing to re-open the famous Zajonc-Lazarus debate), this assumption "led to a series of misrepresentations of appraisal theory. In fact, however, Lazarus himself and essentially every other appraisal theorist after him has pointed out that appraisals can be and often are automatic and outside of awareness (Kappas, 2006). Thus it is entirely
conceivable that 'challenge' appraisal have an implicit underpinning (Frijda, 1986; Smith & Kirby, 2001).

However, a fundamental difference between 'challenge' and 'threat' appraisals is the notion of expected outcome. Whilst 'threats' are associated simply with a negative outcome, 'challenges' encompass the notion of controllability of outcome (Kausar and Powell, 1999), which is intricately linked with the individual's self-evaluation of his/her coping strategies and coping resources (Lazarus, 1999). Thus 'challenge' appraisals occur as a consequence of the individual's perception of their ability to manage the demands of the stressor (Folkman & Lazarus, 1985), whereas 'threats' centre on the potential harms and losses concerned with exposure to the stressor (Lazarus & Folkman, 1984). The key difference between the two processes, therefore, is the emphasis on coping in 'challenge' appraisals.

Coping encompasses the "conscious volitional efforts to regulate emotion, cognition, behaviour, physiology, and the environment in response to stressful events or circumstances" (Compas et al., 2001, pg89). Although this is a vast and diverse literature (see Skinner et al., 2003 for a review of over 300 different types of coping), a significant body of evidence supports the notion of a dual process system of stress responses (e.g. Compas, 2004; Compas et al., 2001; see also Eisenberg et al., 1997; Skinner & Zimmer-Gembeck, 2007. Such systems identify that:

"Two fundamental processes are involved in self-regulation in response to stress. First, there is a set of automatic processes that are activated in response to stress that are related to but distinct from coping... Second, individuals initiate a
Thus coping, as defined by such dual processes theories, is a purely cognitive process, with close similarity to (Barrett & Campos 1991), and/or identical features of (Bridges & Grodnick 1995) the systems that regulate of cognitive emotion. Indeed, there is strong evidence from neuroimaging studies to support a purely explicit role for coping, as separate coping styles are associated with specific patterns of cortical activity (Stiller et al., 1997; Sander et al., 2002) and a very large body of research identifies the PFC and Cingulate Cortex as central to coping (see Gianaros & O’Connor, 2011 for a review). Therefore, as ‘challenge’ appraisals are intricately linked with coping, and coping is a cognitive process, we should not expect any concordance between the ‘challenge’ sub-scale of the SAM and the SA-IAT, which is what I found. In effect, this validates the SA-IAT as an implicit measure of stressor appraisal by the process of discriminant technique (Nosek & Smyth, 2007; Nosek & Hanson, 2008).

‘Threat’ Appraisals. Figure A3 demonstrates an apparent correlation between participants’ average scores on the ‘threat’ sub-scale of the SAM and their SA-IAT D score. In keeping with guidance given in Clegg (1990) for analysis of paired data, I conducted a statistical analysis of correlation using the Spearman’s Rho Correlation Coefficient for ordinal data. The test produces an r value of +0.331, which is indicated by the line of fit shown in Figure A2. This constitutes a relatively weak correlation between ‘threat’ score and SA-IAT D score. Indeed, as the
critical value for this test (p=0.05) is 0.364 the correlation must be viewed as insignificant.

However, looking at Table A5, it becomes apparent that one of the items in the ‘threat’ sub-scale of the SAM (Question 11) has limited the reliability of the sub-scale to some degree (the low ‘corrected item-total correlation’ value and the increased a score if deleted support this conclusion). If I exclude this question from the data and re-calculate the statistics, the Spearman’s Rho Correlation Coefficient becomes 0.467, making the correlation significant at both the 5% (\(r_{\text{crit}} = 0.364\)) and 2% significance level (\(r_{\text{crit}} = 0.432\)). This correlation is presented in Figure 5 below.

![Figure A5](attachment:figure.png)

**Figure A5:** Graph showing the correlation between average score on the adjusted ‘threat’ sub-scale of the SAM against participants’ SA-IAT scores.

From this adjusted data I draw the conclusion that ‘threat’ appraisals and SA-IAT measures demonstrate concordance in that they both measure the “‘fast, reactive, emotionally driven, impulsive ‘hot’ system that
appraises and reacts to external stressor stimuli or situations relatively automatically and with little [immediate] conscious control" (Skinner & Zimmer-Gembeck, 2007). This 'hot' system is automatic, in that it occurs “outside of an individual's awareness” (Compas, 2006, pg227) and, as such, utilises subcortical emotion systems, particularly the amygdala (Davis & Whalen, 2001). In contrast to the volitional, coping-based system, the implicit preconscious system is responsible for short-term adaptive responses, which "bring the organism into a state of readiness to act in accordance with the emotional urge, whether that be to flee, protest, or approach" (Compas et al., 1997, pg110). As 'threat' appraisals exist for this specific purpose (i.e. responding rapidly to the potential negative impact of a stressor) it has been suggested that they make more significant use of the preconscious system that forms the underlying foundation of dual process theories of stress response (Compas, 2004; Compas et al., 2001). Therefore, it follows that 'threat' appraisals and SA-IAT responses both measure the automatic, sub-cortical (and thus exclusively implicit) 'hot' appraisal system that underpins the stress response system (Compas, 2006) and thus, by demonstrating significant concordance, the adjusted 'threat' sub-scale of the SAM validates the use of the SA-IAT as a measure of implicit stressor appraisal.

**Conclusion of IFS Study**

I conclude that the SA-IAT constitutes a valid measure of the implicit appraisal of stressor stimuli.

**Limitations of IFS Study**
Despite drawing the conclusion above, there are a number of limitations to this study, which detract from the validity of the conclusion as presented. I list these limitations as follows:

1. The SA-IAT measures the relative strength of a pair of associations, rather than the absolute strength of single associations (Greenwald & Farnham, 2000). The consequences of this limitation are significant and, as yet, not fully explored in the literature. For example, I cannot assume that relative associations between one pair of semantic categories generalise to other categories, even if they appear similar. For example, the SA-IAT pairings between 'stressful' and 'relaxing' categories would not necessarily yield similar results for the categories 'worrying' and 'calm', despite the similarity of the concepts. This is because the very nature of the pre-attentive or 'automatic evaluation' process (Greenwald & Farnham, 2000) is based on experiential (and, therefore, highly subjective and contextual) associations between explicit stimuli per se. Without the accompanying cognitive appraisal process (which would identify 'relaxing' and 'calm' as being synonymous) I cannot assume that one person’s 'relaxing' is in any way similar to another person’s 'calm'. The same criticism can also levelled at individual SA-IAT scores, where I cannot conclude that participants with identical SA-IAT scores evaluate individual concepts in the same manner. Instead, I can only deduce that the weighted evaluation between shared concept pairs has elicited a similar degree of relative association.
Thus the data produced by the SA-IAT must be taken as having ordinal value, which limits its statistical power considerably and is the reason for using the Spearman's Rho statistical test in place of the Pearson Product Moment Correlation Coefficient.

2. Following on from the previous point, I have already set out Blanton and Jaccard’s (2006) argument that IATs should be considered to be arbitrary instruments on the grounds that they utilise relative scales, rather than normalised values. However, although I have endorsed Nosek and Sriram's (2007) reasoning for such a normalisation process being inapplicable to IAT instruments, nevertheless I accept Blanton and Jaccard’s reservations as applied to this study and, therefore, given the relative nature of the SA-IAT scores recorded in this study, I caution readers that the extent to which my conclusion generalises to other research may be limited.

3. As discussed in the validity section of this project, individual differences between participants have been shown to affect associations in some IAT studies. Although, to a degree, I have controlled for the effect of age (Mellott & Greenwald, 2000) and gender (Greenwald & Farnham, 2000), I have not taken measures to consider the effect of racial, ethnic or cultural differences (Greenwald et al., 1998) on IAT effects. Although many of these individual differences will be limited by choosing participants from the same school community, it is inevitable that individual differences between participants will affect the stimulus appraisal process to some degree.
4. Whilst I have justified the use of the SAM to assess implicit appraisal processes, nevertheless some researchers assert that explicit measures have a limited role in this kind of research (e.g. Shields & Steinke, 2003). For example, in a discussion of the limitations of self-reports in appraisal studies, Kappas states that:

"self-report measures are bound to produce at best a mix of recalled reflective appraisals and reconstructed appraisals that have little to do with what happened in the participants' brains during the recalled event" (Kappas, 2006, pg968)

Accepting this conclusion would undermine the use of the SAM in this type of validation study.

5. As discussed above, familiarity of stimulus items may well have confounded, or even been responsible for, the IAT effects observed in this study (Rothermund & Wentura, 2004). Whilst I have presented evidence that undermines this conclusion (e.g. Dasgupta et al., 2000, Ottaway et al., 2001), I have not attempted to control for stimulus asymmetry in this research and, therefore, I accept that discrepancies between the familiarity of stimulus items across participants may well limit the validity of my conclusion.
Appendix C

Permission to use GSES
Permission granted

to use the General Self-Efficacy Scale for non-commercial research and development purposes. The scale may be shortened and/or modified to meet the particular requirements of the research context.

http://userpage.fu-berlin.de/~health/selfeval.htm

You may print an unlimited number of copies on paper for distribution to research participants. Or the scale may be used in online survey research if the user group is limited to certified users who enter the website with a password.

There is no permission to publish the scale in the Internet, or to print it in publications (except 1 sample item).

The source needs to be cited, the URL mentioned above as well as the book publication:


Professor Dr. Ralf Schwarzer
www.ralfschwarzer.de
Appendix D

The Examination Self-Efficacy Scale (ESES)
Participant ESES Questionnaire

DO NOT PUT YOUR NAME ON THIS FORM

Please enter the first three letters of your mother’s maiden name and the last two digits of your phone number (For example John Doe 078948465 would enter SMI65 as his mother was Jane Smith!)

For each of the following statements, please tick the choice that is closest to how true you think it is for you. The questions ask about your opinion. There are no right or wrong answers.

1. I can always manage to solve difficult exam questions if I try hard enough.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. I do not know what I want to do at university.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. It is easy for me to complete examination questions accurately.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. I am confident that I could deal efficiently with unexpected exam questions

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. In an academic exam, it is important to say what you think

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. I am not good at communicating my point in exams

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Having people feel sorry for me makes no difference in performing well in an exam.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Its important that other people know I’m right.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9. When I am confronted with a problem in an exam, I can usually find the solution.

10. It doesn't matter if you get something wrong in an exam, if you learn from it.

11. I think the results can sometimes justify the means.

12. I spend time identifying my academic goals.

13. I think about my academic future a lot.
14. When I need to think I spend time by myself.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

15. I feel responsible for my academic future.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

16. I feel driven by my personal values to succeed in exams.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17. My parents are a source of academic inspiration.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

18. Little of my success in exams happens because I am lucky.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

19. My friends help me do well in exams.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
20. There are few academic challenges that I cannot control if I focus my efforts on them.

<table>
<thead>
<tr>
<th>Not at all true</th>
<th>Hardly true</th>
<th>Occasionally true</th>
<th>Moderately true</th>
<th>Mostly true</th>
<th>Exactly true</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Appendix E

The 10-item Perceived Stress Scale (PSS10)
Perceived Stress Scale

DO NOT PUT YOUR NAME ON THIS FORM

Instead, please enter the first three letters of your mother’s maiden name and the last two digits of your phone number (For example John Doe 078948465 would enter SMI65 as his mother was Jane Smith!)

........................................................................................................................................................................................................

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate by ticking how often you felt or thought a certain way. There are no right or wrong answers. Please tick one box only per question.

1. In the last month, how often have you been upset because of something that happened unexpectedly?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. In the last month, how often have you felt that you were unable to control the important things in your life?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. In the last month, how often have you felt nervous and "stressed"?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. In the last month, how often have you felt confident about your ability to handle your personal problems?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. In the last month, how often have you felt that things were going your way?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. In the last month, how often have you found that you could not cope with all the things that you had to do?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. In the last month, how often have you been able to control irritations in your life?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
8. In the last month, how often have you felt that you were on top of things?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. In the last month, how often have you been angered because of things that were outside of your control?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

<table>
<thead>
<tr>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Fairly Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

End of Questionnaire!

Many thanks for participating in this study. TWF.
Appendix F

Confirmation letter from the Ethical Research Ethics Committee
Dear Tim

RE08-12 – Buffering preconscious stressor appraisal: the protective role of Self-Efficacy.

I am writing to confirm the Research Ethics Committee of the School of Sport and Education received your application connected to the above mentioned research study. Your application has been independently reviewed to ensure it complies with the University/School Research Ethics requirements and guidelines.

The Chair, acting under delegated authority, is satisfied with the decision reached by the independent reviewers and is pleased to confirm there is no objection on ethical grounds to grant ethics approval to the proposed study.

Any changes to the protocol contained within your application and any unforeseen ethical issues which arise during the conduct of your study must be notified to the Research Ethics Committee.

On behalf of the Research Ethics Committee for the School of Sport and Education, I wish you every success with your study.

Yours sincerely,

Dr Richard Godfrey  
Chair of Research Ethics Committee  
School Of Sport and Education

397
Appendix G

Participant Briefing
Before any data collection began the following statement was read to the participant. After which the participant was asked if she had understood the briefing. If she agreed that she had, she was given an opportunity either to ask questions, to withdraw from the study, or to read and sign the consent form.

Thank you very much for volunteering to participate in this study! Your time is greatly appreciated and will really help with research on how different people appraise different stimuli.

The study consists of two short phases:

1. An initial session in which you will be asked to complete two short 10min questionnaires. Each question in the questionnaire has either five or six answer categories. Please tick just one of the categories, the category that you feel best reflects you.

2. After you have completed the questionnaires we would like you to take a brief 5min word-association task on the computer. Instructions for completing the task are given on the computer at the start of the task.
Any data collected in this study will be kept entirely anonymous. If you want to leave the room at any point and withdraw from the study feel free - you can do that without any repercussions at all. Also, should you wish it, you can withdraw your results from the study after you’ve finished.

If you’re still happy to proceed, please read and sign the consent form and I’ll answer any questions you may have.

Thank you very much!

TWF
Appendix H

Participant Consent Form
Participant Consent Form

I hereby give my consent for my participation in this project.

I understand that the objective of this study is to study certain aspects of how examinations affect Sixth Formers.

I understand that for this study I will be asked to fill out two short questionnaires and complete a brief computer test. I am happy to do this.

I understand that the person responsible for this study is Mr. Filtness. I understand that Mr. Filtness’ supervisor, Professor Watts, has agreed to answer any inquiries that I may have concerning the procedure of this study. I have been informed that I may contact Professor Watts at the School of Sport and Education at Brunel University if I have further questions. Professor Watts’ email address is Mike.Watts@Brunel.ac.uk.

I understand that this study is not dangerous to my health. I also understand that all information concerning this experiment will be coded and that my name or other identifying information will not be used in a way that will link me as a subject in the experiment. I understand that I will be given more details about the study after I have finished participating in it. I understand that I may discontinue my participation in this study at any time I choose.

Signature of participant: Counter-signature of parent: 
Date: Date: