

Self-Evaluation in an Experience of Shame

Akosua Bonsu

UCL MPHIL.STUD THESIS 2011

Thanks to

I would like to thank the members of the philosophy department at University College London, in particular, Dr. Lucy O'Brien who supervised the thesis. I owe a debt of gratitude to graduate students at UCL for many helpful discussions, particularly Craig French.

Contents

Introduction

1. The Dominant View p. 4
2. Self-Evaluations p. 7
3. Breakdown of Chapters p. 14

Chapter 1

1. Cognitive Theories of Emotions in Psychology p. 17
2. The Danger of Equivocation p. 21
3. Empirical Support for Appraisal Theory p. 25
4. Methodological Worries p. 29
5. Chapter Summary p. 37

Chapter 2

1. Weak and Strong Versions of The Dominant View p. 39
2. Recalcitrant Shame p. 43
3. Shame by Association & Vicarious Shame p. 48
4. Chapter Summary p. 51

Chapter 3

1. Hume's Psychological Account of Shame p. 53
2. The Moral to Draw from Hume's Theory of Pride & Shame p. 57
3. A Special Look at Hume on Vicarious Shame p. 63
4. Chapter Summary p. 66

- Conclusion** p. 68

Introduction

1. The Dominant View

The purpose of this essay is to argue against the claim that shame is, or essentially involves, a negative self-evaluation. I refer to this claim about shame as 'The Dominant View' (DV). Within the range of cognitive theories of emotions, DV seems to have acquired the status of a platitude.¹ Gabrielle Taylor (1985 p.1), for example, introduces her seminal book on the emotions with the claim that "[a belief that one] has deviated from some norm" is constitutive of shame, pride and guilt.² Because these beliefs "amount to an assessment of the self" she labels the experiential episodes picked out by the concepts *shame*, *pride* and *guilt* *emotions of self-assessment*. Taylor goes on to claim that shame has two distinguishing features, one of which is "the self directed *adverse* judgment of the person feeling shame." This judgment, we are later told, is "constitutive of the emotion."³ In sum, for Taylor, shame is an emotion which *constitutively* involves a negative self-evaluation. Likewise, Robert Solomon (1993 p. 187) expresses a commitment to DV when he writes "my shame *is* my judgement to the effect that I am responsible for an untoward situation or incident." By identifying shame with a negative self-evaluation Solomon's position is stronger than that of Taylor's in this respect: while Taylor allows for the possibility that shame may have other constitutive features Solomon does not. Solomon's position is unusual because of its strength.

The more usual expression of a commitment to DV is found in the claim that, for adult humans, shame *always involves* (in a way cashed out by the theory) a negative evaluation of the shamed subject. For example, Michael Lewis (1992 p. 75) writes that shame is "a consequence of a failure evaluation." Or more exactly, (Lewis 2002 p. 1181) "to be in a state of shame I must compare my actions against some standard, either my own or someone else's. My failure, relative to the standard, results in a state of shame." For Lewis then, the way in which the involvement of a negative self-evaluation is cashed out in his account of shame is *causal*. The thought is that shame always involves a negative self-evaluation because an adult is caused to be in a shame state only when she measures her actions in relation to some standard, value, or goals and finds that she falls short of them. Here two features are stressed. Firstly there is the *measuring*. Secondly there is the judgment that one has fallen short of that measure: a judgment which may express itself in the "global punitive judgment" that one is a failure *or* more simply in a subjects

¹ A cognitive theory of emotions is an account of the emotions in which propositionally structured attitudes such as beliefs or judgments play an essential role.

² Taylor 1985 p.1

³ Taylor 1985 p. 64

acknowledging that she has failed according to some criteria. These different expressions allow for different phenomenal characterisations of the psychological state of the shamed subject. If shame merely requires the judgment that a subject has failed according to some criteria then it *need not* involve the sort of self loathing with which it is typically associated. It is perfectly consistent that I maintain high self esteem at precisely those moments when I acknowledge that there are criteria that I fall short of. If however, shame requires the judgment that I am a failure, then a phenomenal characterisation of a shame state which excludes a sense of self loathing will not be possible. It is almost invariably assumed that a judgement of the form “I am a failure”, that is a global punitive judgment, is required to experience shame. For example, Deonna and Teroni (2012 p. 72) claim that “shame involves an all-encompassing evaluation of the self as globally unworthy or degraded.”⁴ Similarly, Tagney and Dearing (2003 p. 2) write that shame and guilt are emotions where “in the face of transgression and error, the self turns towards itself, evaluating and rendering judgement.” Shame, on this view “is an emotion of self-blame, involving negative evaluations of the global self.” Also representative of this sort of view is H. B. Lewis (1987 p. 107) who writes “the experience of shame is directly about the *self*, which is the focus of a negative evaluation.” Presumably the popularity of the claim that shame requires a global punitive judgment is to cement the connection between shame and self loathing or a lowering of one’s self esteem? Whatever the case may be, the idea that shame invariably involves a judgment that I am a failure will be objected to in this essay. I will also object to the idea that shame invariably involves an acknowledgement of the judgment that I have failed according to criteria.

So far three types of relations between the experiential state of shame and a negative self-evaluation have been considered indiscriminately. These are the relations of constitution, the relation of cause and the relation of identity. To count as a proponent of DV all that matters is that some relationship is taken by a theorist to be, in some sense, required to experience shame. Different versions of DV can be identified by the relationship that they claim exists between shame and a negative self-evaluation. It is my view that DV is false: there need be no relation between the state of shame and a negative self-evaluation in order for a subject to experience shame. Although I believe that collectively my arguments against DV apply regardless of what specific claim about the relationship between shame and a negative self-evaluation is made by its proponent, my arguments will be specifically directed at a version of DV which posits a *causal* relationship between shame and a negative self-evaluation. More exactly, the version of DV argued against here is the one proposed by an Appraisal Theory (AT).

⁴ For other proponents of DV c.f. Velleman 2006 p. 46-7 fn 1

AT represents itself as an *empirical theory* of emotional elicitation, which is to say, the appraisal theorist believes that the claims he makes about emotional elicitation can be empirically supported. We can divide emotional types into two categories: *basic emotions* and *self-conscious emotions*. Basic emotions include fear, anger and disgust. Self-conscious emotions include shame, guilt, jealousy, love and pride. The distinguishing features of basic and self-conscious emotions are a matter of debate. Most theorists accept that basic emotions are associated with universal *facial* expressions and that self-conscious emotions are not.⁵ The key difference between basic and self-conscious emotions, for an appraisal theorist, is that the elicitation of a self-conscious emotion is mediated by cognitive processes which involve or result in self-evaluations.⁶ The appraisal theorist will agree that particular instances of fear, anger and joy *may* also be elicited by cognitive processes involving self-evaluations but he will insist that these basic emotions do not count as self-conscious because, in general, these emotions can be elicited by processes which do not involve self-evaluations. In sum, an appraisal theorist accepts the claim that:

- (1) When a stimulus *x* elicits a self-conscious emotion, *x* prompts the subject to engage in a pattern of thinking which involves or results in a self-evaluation.

When attention has turned to shame in particular, the appraisal theorist makes the further claim that shame is caused by an event *x* only if *x* prompts a subject to make a negative self-evaluation. Thus, the version of DV advocated by an appraisal theorist and objected to in this essay is this:

- (2) When a stimulus *x* elicits a shame response, *x* prompts the subject to engage in a pattern of thinking which involves or results in a negative self-evaluation.

Both (1) and (2) are taken by the appraisal theorist to be empirical claims, which is to say that they are claims that may be confirmed or disconfirmed by experience. Supposing that they are confirmed by experience, however, it is not always clear what status these claims are supposed to have. Are we to read them as statements which are (in some sense) *necessary*: so that a subject *could not* experience a self-conscious emotion absent a self-evaluation? Or are we to read them as statements describing regularities or uniformities, which may, in abnormal circumstances, admit of exceptions? The first reading is suggested by Tracy and Robins (2004 p. 106) who write:

“Self-conscious emotions differ from basic emotions because they require self-awareness and self-representations. Although basic emotions such as fear and

⁵ It is widely accepted that basic emotions are pan-culturally shared emotions that have universal facial expressions. C.f. Ekman 1990, 1992a, 1992b

⁶ Tracy & Robins 2004; M. Lewis 2003a

sadness *can* and often do involve self-evaluative processes, only self-conscious emotions *must* involve these processes... A sense of self, as conceived by self theorists since James (1890), includes an ongoing sense of self-awareness (the "I" self) and the capacity for complex self-representations (the "me" self, or the mental representations that constitute one's identity). *Together, these self-processes make it possible for self-evaluations, and therefore self-conscious emotions, to occur.*" (Italics added)

We have here two claims which suggest that (1) and (2) should be given the necessity reading. First, there is the idea that self-evaluations are required to experience self-conscious emotions. Second there is the modal way of distinguishing between self-conscious and basic emotions. The weaker view that (1) and (2) are statements which express regularities which may, in special cases, admit of some exceptions is suggested in the following passage by Scherer (2001 p. 372-373). Scherer writes:

"[F]ew, if any, appraisal theorists claim that all incidences of affective responses are necessarily produced by appraisal [i.e., evaluative component] or can only be explained by such principles... Thus, if there were a consensus that the physiological, expressive, and possibly experiential patterns produced by electrical brain stimulation, hormonal imbalance, social contagion, or a host of other factors... are to be counted among the real emotion... appraisal theorist are unlikely to worry."

This is because, according to Scherer:

"[t]he percentage of emotion episodes where there is absolutely no trace of appraisal (on at least one of the levels of processing assumed to operate by most theorists) is relatively low... The tendency to focus on emotion episodes that do not seem to have situational causes [i.e., that are not caused by specific evaluations] generates the risk of using rare events as the basis of a theory for the normal case"

These passages suggest that (1) and (2) should be taken to be descriptions of regularities in experience which may, in some cases, admit of exceptions. Importantly, Scherer makes the further point here that those cases in which regularities do not hold are relatively low in frequency, and as such, the appraisal theorists account of emotional elicitation (an account which consists of statements like (1) and (2)) should not be abandoned to accommodate them. Whether or not we should view (1) and (2) as statements describing regular connections or as statements describing a necessary connection is, unfortunately, beyond the scope of this essay. The arguments to follow apply directly to those appraisal

theorists who, like Tracy and Robin's seem to think that certain appraisals (i.e., evaluations) are *necessary* for emotional elicitation. For those like Scherer who take AT to be claiming only that certain emotions are regularly connected with certain appraisals, the arguments will have a different effect. In this latter case, rather than showing why AT is false, the arguments presented here merely force the appraisal theorist to see that in the case of shame, it is not acceptable to sideline those cases in which, I will argue, a negative self-evaluation is absent. Instead, a proper understanding of these abnormal cases should have a direct effect on our understanding of the shame state.

Having now introduced DV in a sketchy form and said something about the version of DV which will be objected to in the essay, the purpose of this introduction is now twofold. First, I wish to get clear on a concept that is central to all versions of DV, namely, that of a self-evaluation. What is a self-evaluation? What is a *negative* self-evaluation? And what capacities must a subject have in order to make a self-evaluation? Second, I wish to give a breakdown of the chapters. In doing so, I aim to get clear on the structure of the thesis and in particular, to show how each chapter contributes to a critique of claim (2) above.

2. Self-Evaluations

A self-evaluation combines two things: the general notion of evaluation and the structural feature of reflexivity. Let's start with the general notion of an evaluation. An evaluation can be viewed either as a process or as an achievement. As a process, an evaluation is a consideration of whether or not an object has a certain value. The value itself, *v*, will set a standard in the light of which a subject can determine whether an object has or lacks *v*. Values set standards in this way simply because value terms have meanings and the meaning of a concept sets conditions which determine when that concept is correctly applied. Admittedly, value concepts are vague concepts. That is to say, their application conditions leave room for borderline cases. But in a lot of cases, the meaning of *good* can determine whether or not an object that has just *these* properties can be classified as good. When an evaluation is considered as a process it simply *is* a subject deliberating about the properties of some object, *o*, accessing her store of knowledge about the application condition of a certain value concept, *v*, and weighing up whether or not *o*, given *o*'s various properties, exemplifies *v*. The result of this evaluative process is an *evaluative judgment* which states whether or not *o* exemplifies *v*.⁷ The result of an evaluative process is not automatically to be thought of as knowledge because there are the following margins of

⁷ The result of an evaluative process may also be a judgment or an opinion about the worth or significance of *o*. That judgment requires more than just knowledge of the application conditions of *v*. It also requires knowledge of the relative significance of *v*.

error: (i) a subject could be mistaken about the properties of *o*, (ii) a subject may know enough about the application conditions of a concept to be classed as knowing its meaning while remaining ignorant of some of the cases in which it would be appropriate to apply that concept. That being the case a subject may fail to see that *o* exemplifies *v* even when she knows the meaning of *v* and she is not mistaken about the properties of *o*. Finally an evaluative judgment will not count as knowledge if (iii) something has gone wrong in the deliberative process. Absent these errors, the evaluative judgment which is the end product of an evaluative process counts as knowledge.⁸

For the majority of this discussion reference to an “evaluation” is reference to an evaluative judgment or an opinion reached on the basis of an evaluative process. Context will help disambiguate those rare circumstances in which I use the term “evaluation” to refer to a process. One might well wonder *why are we talking about evaluations in connection with emotions anyway?* The requirement for a discussion of the notion of an evaluation is said to spring from the intentionality of emotions. Emotions are intentional in the sense that they are directed towards objects. I am not just afraid, but afraid *of* that dog or *of speaking in public*; I am not just ashamed but ashamed *of* my behaviour or *of* my country. The object of an emotion is the thing that you are angry or sad, overjoyed or ashamed, embarrassed or curious *about*. To illustrate the need to invoke evaluations in accounting for emotional experience, let us suppose you are walking home alone at night. At the end of the road, strolling towards you is a fox. If you are afraid at the sight of the fox, then the fox is the object of your emotion. Even if, on closer inspection, no fox is present, the fox *was* the object of your fear. Evaluations are introduced into the picture once we recognise that it is not the fox *per se* that makes you afraid. If that were the case, everyone who encountered that object in the circumstances described, would be afraid of it, and that isn’t true. What generates fear of the fox here is the significance that that object has for you. In other words, it is because you think of the fox approaching as a bad thing (perhaps because it constitutes a threat or danger to you) that you feel the way you do. So the thought goes, the experience of an emotion requires a subject not only to represent its’ object but to represent its’ object as having some evaluative property. The representation of the object as having an evaluative property is what makes the subjects’ reaction to it an emotional one rather than simply a matter of mere cognition. This is because a subjects’ evaluation of the object represents the significance it has for *her* and it is the objects’ subjective significance that makes her reaction to it an emotional one. In the simplest case, a subject represents the object as being good or bad. It appears to be the fact that a subject represents the object in this way that puts her in the particular emotional state that she is

⁸ Konzelmann Zic 2011 p. 11-19

in. To return to our example, it appears to be the fact that Tony represents the fox as dangerous (where danger is classified by Tony as bad) whereas Claire represents it as benign and friendly (where benign and friendly is classified by Claire as good) that Tony experiences fear on sight of the fox and Claire experiences, let us suppose, curiosity.

In sum, evaluations come into the picture because of the unique way in which emotions are intentional. Emotions are uniquely intentional because a subject must not only represent the object of an emotion but she must represent it (at the very least) as being good or bad. How a subject represents the emotions object is determined by the significance of that object *for her*. Note that the significance of the emotions object will not only depend on the objective properties of that object, but also on subjective variables such as the subjects own history, her culture and her mood. In fact, it is because emotional evaluations involve subjective variables that we think of emotions, colloquially as subjective reactions to their objects. More precisely, it is the fact that emotions involve subjective evaluations of their objects that different people often have different emotional reactions to the same object and the same person can have different emotional reactions to the same object at different times.⁹

So much for the notion of an evaluation and why it is that that notion should play a role in an account of the emotions. Our next question must be, what does the structural feature of reflexivity add to the notion of an evaluation? That is to say, what is a self-evaluation? Very simply, a self-evaluation is an evaluative judgment about the self made by the self. What happens when a subject makes a self-evaluation is that she engages in a deliberative process that results in the judgment that *I am x* where *x* is an evaluative predicate such as “good” or “bad”. To make a self-evaluation a subject must *at least* have (a) a conception of herself, i.e., a representation of some of her characteristics (b) the ability to focus attention on that self-concept *in the knowledge that it is a representation of herself* (the apprehending subject) that she is focusing attention on and finally (c) knowledge of the application conditions of a value, *v*, and as such criteria on the basis of which she can determine whether, given the characteristics represented in her self-concept, she has or lacks *v*. It’s important to be clear on these three elements.

A self-concept is a representation that the subject has of the characteristics that she believes herself to possess. However, not all of a subject’s characteristics will enter into her self-concept. For example, a self-concept will not, at least not typically, include relational properties such as sitting to the left of the lamp post. I take it that a self-concept includes only those features which are believed to be relatively stable characteristics of

⁹ Sroufe 1995; Lazarus 1991; Frijda 1988

the subject and which a subject cares about possessing.¹⁰ A subjects' self concept therefore corresponds to some amongst the beliefs that she has about herself at a given time. For example, my own self-concept includes my values, some of my physical properties, my achievements, my goals, my ethnicity, my working class background and so on. In virtue of representing these aspects of a subjects' character, we can take it that a self-concept underwrites the idea that a subject has of herself as a *unique* individual. Henceforth I will label a self-concept, as I have just characterised it *my idea of myself*. It is important to distinguish *my idea of myself* from *my idea of myself for others*. My idea of myself for others is a representation of those characteristics which I care about possessing (or lacking) but which I believe others attribute to me. That is to say, it is a representation of how I believe I am represented in the minds of others. Of course, if I judge that those others represent me accurately, my idea of myself for others collapses into my idea of myself. However it is important to keep these two ideas of the self separate as they can, and often do, come apart. It is true that both my idea of myself and my idea of myself for others are a sort of self-conception because both are representations of me, however only the former here is labelled a self-concept because only the former underwrites a subjects own conception of herself as a unique individual. Before moving on to the notion of self-directed attention, we should note a further difference between my idea of myself and my idea of myself for others. This is a difference in the complexity of capacities required to have an idea of myself and an idea of myself for others. Having an idea of myself presupposes only that I have an ability to characterize myself in abstract ways (i.e., the ability to self-represent). Having an idea of myself for others requires that I have an ability to represent myself as I might be represented in the mind of others. The latter is a more complex ability because it requires *thinking about myself from the perspective of another*. To make a self-evaluation either my idea of myself or my idea of myself for others is required because either can act as the object of the evaluation.

The ability to direct attention on the self is also required to make a self-evaluation. To exercise this ability, it is required that a subject has an ongoing sense of herself as a subject of experience (i.e. as the thing which attends), an idea of herself or an idea of herself for others (i.e., a representation of the self which acts as the object onto which attention is directed) and finally, an awareness that the representation which she directs attention onto *is* a representation of herself. This final feature is essential to distinguishing a self-evaluation from a thought a subject has which happens to be about herself. Suppose, for example, that Charlie unknowingly enters a hall of mirrors of various shapes and sizes which distort the reflected image. On sight of the mirror image, he may well think, *that*

¹⁰ Frankfurt 1988 p. 80-95

person is beautiful. As beauty is a value and the demonstrative pronoun refers to Charlie, Charlie has here succeeded in making an evaluative judgment about himself. This evaluative judgment, however, falls short of a *self-evaluation* because the demonstrative pronoun (*that person*) could not be replaced with the first person pronoun (*I*). Charlie would not, let us suppose, assent to the judgment “I am beautiful” even though he does assent to the judgment “that person is beautiful.” Self-evaluations always take the form *I am/am not such and such* and one of the many peculiarities of the first person pronoun is that one cannot use it without an awareness that it refers to oneself. So in order to make a self-evaluation, an evaluative judgment of the form “I am/am not such and such”, a subject must direct attention on a self representation in the knowledge that the object evaluated and the evaluator are one and the same.¹¹

It is sometimes said that the sort of self-awareness required to experience shame is not achieved by a subject who merely exercises a capacity to think about herself in *any* way. According to these theorists, a subject who experiences shame must think about herself in a particular way, specifically, she must think about herself from the perspective of another. This idea, endorsed mainly by theorists who think that shame is essentially related to an audience, can be simply put as follows: the self-awareness involved in experiencing shame requires a subject to direct attention on her idea of herself for others. More specifically, the idea is that the self-awareness presupposed in making the self-evaluations required to experience shame is *objective* in this sense: the subject who makes this evaluation must think about herself from the perspective of another. So according to these theorists, in experiencing shame, a subject not only represents but evaluates her idea of herself for others.¹² The difference between the type of account in which the centrality of an audience is reflected in the sort of self-awareness required to experience shame and those accounts which claim that attention directed at the idea of the self may be sufficient to account for an experience of shame is that in the latter cases, the requisite form of self-awareness does not proceed through thoughts about another’s attitudes towards the self. Although this difference is significant it should not be overrated. Both types of account claim that *some* form of awareness that has a self-representation as its object is required to experience shame and both agree that these self-representations are composed of concepts which are socially acquired, and as such these are representations of myself that others *could* have of me.

The final feature identified above as necessary for making a self-evaluation is the possession of criteria. Why are criteria required? Criteria are required because no

¹¹ Baker 2011 p. 34

¹² Sartre 1943/1957; Lewis, M. 1999; Lewis, H. 1971; Taylor, 1985

evaluation of x as y can proceed without evaluative criteria determining what it takes for x to be classed as y . It should be noted that having the criteria employed by the subject in these cases is based partly on an understanding of the linguistic norms which govern the correct use of value terms and partly on an understanding of the relative import of those values. As such, possession of criteria is based on socialisation. By socialization I mean being brought up in a linguistic community. A linguistic community is required to learn the meaning of value terms and therefore to possess the criteria used to assess whether or not a subject possesses a value. A linguistic community is also required to learn the relative importance of values and therefore informs judgments of how much weight certain values should hold when deliberating. It is *partly* for these reasons that proponents of DV have stressed that shame is a social emotion.¹³

The three enablers of a self-evaluation work in the following way: to produce a self-evaluation the subject of the experience reflects on a mental representation of herself in order to determine whether, according to the criteria for the application of a value term v , *she* (the person represented) succeeds or fails in exemplifying v . From this brief summation, we can see that if self-conscious emotions require self-evaluations, the experience of such emotions requires a conceptually sophisticated subject. The subject of a self-conscious emotion must engage in mental processes that depend on the manipulation and storage of representations which are derived from learning or experience.¹⁴ She must be a member of a linguistic community who has grown to understand certain value terms and to have some grasp of their relative importance. A subject who lacks a self-representation, or the ability to reflect on a self representation, or who hasn't acquired the socialisation required to know the relative import of and application conditions of certain value terms is automatically barred from experiencing self-conscious emotions. Such a subject, according to this view, could not experience shame.

So far an attempt has been made to explicate DV by getting clear on the notion of a self-evaluation. What is required now is a way of understanding what it is for a self-evaluation to be positive or negative. There are at least two ways in which we can understand the notions of a positive and a negative self-evaluation. We may suppose that (1) a subject makes a *positive* self-evaluation when she judges that she passes the criteria on the basis of which she would count as having an evaluative property and that she makes a *negative* self-evaluation when she judges that she fails that criteria *or* we may suppose that (2) a subject makes a positive self-evaluation when she judges that she is x where x is a positive value such as "good", "intelligent" or "kind" and that she makes a *negative* self-evaluation

¹³ Deonna & Teroni 2011b

¹⁴ Izzard 1993

when she judges that she is x where x is a negative value such as “bad”, “stupid” or “unkind.” On the first reading a subject makes a negative self-evaluation when she judges “I have failed to exemplify an evaluative property based on these criteria.” On the second reading a subject makes a negative self-evaluation when she judges “I am a failure”, “I am unworthy”, “I am a loser” and so on. This second way of understanding a negative self-evaluation was referred to above as a global punitive judgment. The difference between a global punitive judgment and the judgment that “I have failed to exemplify a value” is that the former requires a subject to think badly of herself but the latter does not. To judge that I have failed to exemplify a value will not require me to think badly of myself if I do not endorse the criteria on the basis of which that judgment is made. Endorsement is a further step that sometimes connects a global punitive judgment to the judgment that “I have failed according to this criteria.” To illustrate the point, reconsider the passage in which Lewis (2002 p. 1181) writes:

“[T]o be in a state of shame I must compare my actions against some standard, *either my own or someone else’s*. My failure, relative to the standard, results in a state of shame.” (Italics added)

The section italicised points to the two ways of understanding the notion of a negative self-evaluation just outlined and therefore this passage can be used to indicate the two things which we might mean when we say that I am required to make a negative self-evaluation when I experience shame. If the standards against which I am judged negatively are my own, that is to say, if I accept those standards, then I am more likely to endorse a global punitive judgment when I note that I have failed according to those standards. For example, accepting criteria determining that homosexuality is wrong and judging that I am gay, I will also judge either that I am a failure or that I have acted wrongly. Which is to say, I am more likely to make a global punitive judgement if, in shame, I accept the criteria according to which I am negatively assessed. If, however, I do not accept that criteria, it is very unlikely that I will endorse a global punitive judgment: simply judging that I am bad according to *your* standards will not rationally require or compel me to judge either that I am bad or that I have acted wrongly. For example, I can judge that being gay is bad according to your standards, judge that I am gay and fail to think of myself or my behaviour as in any way bad in virtue of my sexual preferences. Although proponents of DV mostly claim that the self-evaluations required to experience shame are global and punitive, I will understand DV as claiming that shame requires a negative self-evaluation either in the sense that it requires a subject to make a global punitive judgment *or* in the sense that it requires the subject to judge that she has failed according to evaluative

criteria she may or may not accept. In what follows counterexamples to both types of view will be presented.

3. Summary of Chapters

The purpose of the thesis is to critique DV (DV) according to which an experience of shame requires a subject to make a negative self-evaluation. Negative self-evaluations are judgments of the form “I am x” where “x” is a negative evaluative predicate or they are judgments of the form “I have failed to exemplify an evaluative property on the basis of some criteria.” To make a negative self-evaluation a subject must have sophisticated cognitive abilities such as the ability to direct attention onto a self representation and an understanding of the application conditions of various value terms. In what follows, the critical arguments will be directed towards the Appraisal Theory(AT). AT is a version of DV that takes the relation between shame and a negative self-evaluation to be causal. I believe that some of the arguments given against AT will, however, give us good reason to abandon DV all together. The thesis is composed of three chapters, which I will now briefly summarize.

In chapter 1, AT is introduced in the context of a discussion on cognitive theories of emotions in contemporary psychology. According to an appraisal theorist, if a stimulus x elicits *shame*, then x prompts the subject to engage in a pattern of thinking which involves or results in a *negative* self-evaluation. AT is particularly interesting in that it attempts to establish claims about emotional elicitation by empirical means. The attempt to devise experiments which support the view that shame requires a subject to make a negative self-evaluation raises a host of methodological worries. I critically discuss whether, given these methodological worries, the experiments conducted by an appraisal theorist can give us more than anecdotal evidence for the connection between shame and negative self-evaluations. I conclude that while there can be empirical support for *some* of the associations between appraisals and emotions, the appeal to this evidence has a limited value in the case of shame. These experiments cannot give us a reason to believe that shame is not, perhaps frequently, experienced when a subject does not make a negative self-evaluation. Indeed, the purpose of chapter 2 is to introduce and explore such cases.

Chapter 2 presents a range of cases in which either (1) a subject experiences shame but fails to make a global punitive judgment or (2) a subject experiences shame but fails to make a self-evaluation of any kind. In the first of these cases, the subjects' shame is recalcitrant. An emotion is recalcitrant if it conflicts with the evaluative judgment with

which it is naturally associated.¹⁵ For example, if I am ashamed of being poor even though being poor is not a state of affairs which I judge negatively, my shame is recalcitrant. This is because being ashamed of *x* is naturally associated with a negative evaluation of *x*. I argue that in recalcitrant shame a subject does not make a global punitive judgment because she does not accept the criteria on the basis of which being poor is bad. What explains the shame response in this case is not the subjects' belief that being poor is bad but rather the fact that she views her poverty from the perspective of another subject who would make that judgment. That is to say, in experiencing recalcitrant shame a subject employs evaluative criteria which she does not herself endorse to judge that she has failed to exemplify value *x*. But why does the awareness that another would judge one's poverty negatively generate a shame response? I take it that what explains the fact that a subject here responds with shame is that the criteria she employs to judge that she has failed is criteria held by some representative member of a group to which she belongs.

Recalcitrant shame does not act as a counter-example to versions of DV which claim that in order to experience shame, it is sufficient for a subject to judge that she has failed according to some criteria. In the second part of chapter 2, I introduce two cases in which, I take it, a subject experiences shame but makes no self-evaluation at all. These cases are vicarious shame and shame by association. A subject is shamed by association when something with which she is associated is negatively evaluated. For example, a subject who is ashamed of her country for invading a sovereign state is ashamed by association. Here the subject does not think of *herself* but rather of the associated object as failing from her own perspective or from the perspective of another. The same point applies to vicarious shame. A subject experiences shame vicariously when she experiences shame as a result of simulating a circumstance in which another subject does something that she takes to be worthy of a shame response. Here it is not the self but an imagined other who is judged negatively. In this case, a subject does not think of *herself* but another (who may or may not be closely related to her) as failing either from her own perspective or from the perspective of another. It seems then that the pattern of thinking which generates shame by association and vicarious shame not only fail to include a negative self-evaluation they fail to include any sort of *self-evaluation* at all. This conclusion suggests that we should not take it as given that in experiencing shame a subjects attention is directed onto a mental representation of *herself* – the type of awareness which does explanatory work in the case of shame may well be a focus on a relation between the self and another related object or a focus on another subject who's perspective one imaginatively adopts.

¹⁵ D'Arms and Jacobson 2003

Given the different forms that a shame experience takes in the three examples just considered, the question arises can anything general be said about the causal mechanisms which generate the shame experience? In the final chapter, I discuss and largely endorse Hume's account of the causes of shame. Hume's approach to the emotions is much like the approach adopted by proponents of DV. He aims to advance our theoretical understanding of what shame is by telling a story about the circumstances in which shame is elicited. What is distinctive about Hume's account, for our purpose, is that the notion of an evaluation plays no essential role in his explanation of shames elicitation and yet with simple modifications to Hume's account, we can explain all of the cases above described. Shame, on Hume's view, is simply a response to those events that prompt a subject to experience a particular kind of displeasure in the idea of self. More specifically the idea is that shame is caused by an event x when the subject believes that x is negative and that x is related to me in some way. The critical point that I make in this final chapter is that an interpretation of Hume's account of shame which reads into it a version of DV would rob it of an ability to explain shame by association, recalcitrant shame and vicarious shame. Such an interpretation should be resisted.

To sum up: chapter 1 introduces a version of DV known as Appraisal Theory and reflects on the question of whether or not empirical support can be accumulated for it. Chapter 2 attempts to undermine DV by highlighting cases in which shame is caused by a pattern of thinking which either fails specifically to include a global punitive judgment or fails quite generally to include a self-evaluation of any sort. Given the heterogenous nature of shame by association, vicarious shame, recalcitrant shame and shame involving global punitive judgments, one might be tempted to think that no general account of the causal mechanisms of shame can be given. I suggest in chapter 3 that we can, by modifying Hume's account of the indirect passions, give a model of shame which can accommodate cases in which shame requires a subject to make a negative self-evaluation of some sort and cases in which it does not.

Chapter 1

1. Cognitive Theories of Emotions in Psychology

Cognitive theories of emotions in psychology are invariably theories of emotional elicitation. The central claim of such theories is that emotional experiences are caused by subjective evaluations of a stimulus. Subjective evaluations are a number of assessments made by the subject of the significance of the stimulus for her well being. Experiments conducted by these theorists are designed to uncover which sets of subjective evaluations elicit which types of emotions. A fairly uncontroversial example is the account given of fear. Fear, it is claimed, is caused by an evaluation of a stimulus (e.g. a fast approaching dog) as a danger to my well being. It is claimed that a subject who did not evaluate a stimulus as dangerous or as constituting a danger will not experience shame as a response to it. So the idea is not only that evaluations intervene between a stimulus and an emotional response but also that these evaluations determine the type of emotion experienced in response to the stimulus. There is some debate about how far a set of experiments can justify the particular association of *this* emotion type with *that* set of subjective evaluations. There is also some debate about how to account for the cognitive processes which underpin and enable these evaluations. These are important differences but what matters more here is what connects this group of theories. What connects all these theories are the claims that (1) an evaluation of a stimulus is what triggers an emotional response, (2) specific evaluations trigger specific emotions and what these specific evaluations are can be uncovered through experimentation and finally (3) absent any evaluation, no emotion will be experienced.

In this chapter, I will discuss claims (1) and (2). It is useful to begin the discussion by selecting a representative of the class of psychological accounts of emotions from the perspective of which we can explicate the claim that emotions are caused by subjective evaluations. This is primarily because different theories have different views of the number and kind of evaluations required to trigger a particular emotional response. These differences are not important for the critical points made in this chapter but they are important when it comes to giving an accurate representation of a theorist who puts forward a psychological account of emotional elicitation. I have chosen here to discuss Appraisal Theory (AT) because it is the most popular of the kinds of psychological accounts under consideration. I have chosen to discuss Lazarus's version of AT because he was the first to conceptualise appraisals as *structured processes* composed of multiple parts or 'dimensions' and it is this conception of appraisals which dominates the

psychological literature.¹⁶ In what follows, I will make three critical points about (1) and (2) but I will address these criticisms directly towards AT. The first critical point is that when we keep the meaning of “evaluation” constant, it cannot be true that both basic and self-conscious emotions are caused by evaluations. The second critical point is that the sorts of experiments which are invoked to establish a causal connection between a set of appraisals and an emotion are either beset with methodological worries or they are the sort of results which can be known on the basis of the meaning of the relevant emotional terms. The final critical point is that, the evidence for the claim that shame is caused by self-evaluations is based on experiments that provide no evidence to suppose that shame is not also caused in situations where no self-evaluation is made. This sets the scene for chapter 2 in which various examples of these neglected cases are discussed.

Let’s begin the discussion with the question: what, exactly, are appraisals? Appraisals are structured processes in which a subject evaluates a stimulus. The process is split into multiple parts. At each stage in the process a subject evaluates the stimulus as it relates to her individual concerns. Appraisals are *relational* in the sense that they are not evaluations that a subject makes of the stimulus tout court but rather evaluations she makes of the stimulus given her individual goals, beliefs, needs, desires, values and so on. The set of concerns (or appraisal dimensions) that an appraisal of the stimulus is conditional on is a matter of debate. Lazarus (1991) originally identified six appraisal dimensions which he organised into two groups: primary appraisals and secondary appraisals.¹⁷ Primary appraisals collectively determine the personal significance of the event while secondary appraisals collectively determine what coping options are available to the subject given the outcome of her primary appraisals.¹⁸ More exactly, primary appraisals consist of three sorts of evaluations which Lazarus refers to as goal relevance, goal congruence, and type of ego-involvement. An evaluation of *goal relevance* determines whether and to what extent an event is relevant to a subjects’ well-being, where “being relevant to well being” means something like having a bearing on personal goals. Goal relevance is the most important of all the primary appraisals because if a subject determines that no goal is at stake then no emotion will be experienced. The thought here seems to be something like this: if I do not evaluate the situation as in any way relevant to my well being then I simply will not care about it and as such I simply will not react emotionally to it. An evaluation of *goal congruence or incongruence* determines whether the stimulus event helps or hinders the subjects’ goals. When an event is evaluated as facilitating or advancing the subjects goals, it is goal congruent and the emotion

¹⁶ Schorr 2001b p. 22-3

¹⁷ Lazarus 1991 p. 133-152

¹⁸ Lazarus 1991 p. 133

experienced in response to the event is likely to be positive. If however the event is evaluated as hindering or thwarting a subject's goals then it is goal incongruent and the resultant emotion is likely to be negative. For example, if I evaluate the situation in which a dog is fast approaching as goal congruent because it advances my goal of being reunited with my favourite pet, I am likely to experience a positive emotion such as joy. If, however I evaluate that same event as one in which my goal of avoiding physical harm is thwarted, I am likely to experience a negative emotion such as fear.¹⁹ The final evaluation involved in primary appraisals is an evaluation of *ego involvement*. This is an evaluation of whether goals relating to a subjects ego identity are at stake. Identity goals are any goals which are relevant to, for example, self-esteem, social esteem, moral values, life goals or other persons and their well-being. Interestingly, Lazarus predicts that the sort of identity goals involved in both shame and anger will be *the desire to preserve or enhance self- or social esteem*.²⁰

As we have seen, one of the central aims of the psychological theories under investigation is to type identify emotions on the basis of the set of subjective evaluations required to elicit them. But so far on Lazarus' picture, both anger and shame are caused by events which are evaluated as goal relevant, goal incongruent and as involving goals that specifically relate to ego identity. As such, we cannot determine whether an event will generate anger or shame on the basis of primary evaluations alone. To determine which of these two emotions are experienced in response to an event we must move beyond evaluations of primary appraisals to see what secondary appraisals a subject makes in response to the event. Secondary appraisals consist of three sorts of evaluations which Lazarus refers to as credit or blame, coping potential and future expectations. Evaluations of *credit or blame* are judgments about what or who is responsible for a harm, threat, challenge, or benefit. Here it should be stressed that although evaluations of credit or blame determine who is responsible for the event these evaluations are not merely attributions of responsibility. Beyond such attributions these evaluations include judgements about whether the event was intentional or unintentional and/or whether the situation is one which could have been prevented and by whom it could have been prevented. For example, a parent responding emotionally to a child who has stolen a large amount of money might attribute blame to the child because she judges that his intentions were malevolent. In judging that his intentions were malevolent, that the event is personally relevant and goal incongruent and finally that it involves some among her identity goals, the parent is likely to experience anger. If, however, we subtract the

¹⁹ Here, positive and negative refer to whether an experience of an emotion is pleasurable or painful.

²⁰ Lazarus 2001 p. 57

mothers evaluation that her child's intention was malevolent, replace it with an evaluation of his intention is benevolent (e.g., if the child stole the money from a company to give back to investors after the company has declared itself bankrupt) it is unlikely that the mother will react with anger and quite likely that she will react with pride. Evaluations of *coping potential* are evaluations determining what prospects there are for enacting behaviours or adopting attitudes that will positively influence the situation. These might include beliefs about things that a subject can do (behaviours she might adopt) to eliminate the harm or threat presented by the situation or to bring to fruition the challenge or benefit presented by the situation. For example, the parent of the thieving child might experience anxiety as well as anger if in addition to judging that the child's intent was malevolent she judges that she does not have the ability to cope with the situation. The final evaluation involved in secondary appraisals is an evaluation of *future expectancy* . That is, an evaluation of whether the event, once completed, will continue (in the future) to help or hinder her goals or well being. For example, the parent of the thieving child might experience sadness if she believes that the situation will never improve (perhaps because she thinks her son is a bad apple and that he cannot help but commit acts of criminality).

Let's draw together the central features of the account which Lazarus offers us. For Lazarus, the subjective evaluations which cause an emotional experience fall into two categories: primary appraisals and secondary appraisals. There are three primary appraisals which can briefly be characterised as answers to the following three questions: has something relevant to my goals occurred? Is it congruent with my goals? And how is my ego involved? There are three secondary appraisals which can be characterised as answers to the following three questions: who deserves credit or blame? What coping options are available? And what can I expect for the future? The answers to these questions determine the type of emotion experienced as a response to the stimulus. Critically, the answers to these six questions are then summarised by *a core relational theme* . A core relational theme is not an explicit evaluation that a subject makes when confronted with a stimulus. Rather it is a way of summing up the evaluations involved in the six appraisal dimensions, or if you prefer, a way of summing up the answers to the six questions just identified. More specifically, we are to think of a core relational theme as capturing "the meaning of the full configuration of appraisal components *in a way that transcends the components' individual meanings* - much in the way a sentence captures a complex idea that goes beyond the meaning of individual words."²¹ Consider, for example, sadness caused by the accidental death of a sibling in a car crash. According to this account

²¹ Smith & Lazarus 1993 p. 237

if the event of the cars crashing is to cause sadness a subject must evaluate it as personally significant (appraisal of goal relevance), as incompatible with her desires (appraisal of goal incongruence) as involving someone she cares about (appraisal of ego involvement), that no one is to blame for the situation (appraisal of no blame) and finally that there is nothing she can do about it (appraisal of low coping potential).²² The combination of these five specific evaluations can be summed up by the core relational theme of sadness which is *an irrevocable loss*. Any event that is evaluated as an irrevocable loss, on this picture, will generate sadness irrespective of any differences between them. That is to say, appraising an event as an irrevocable loss is sufficient for an experience of sadness.²³

2. The Danger of Equivocation

Like all cognitive theories of emotions in the psychological literature, an appraisal theorist claims that a subject experiences an emotion only because she evaluates a stimulus in specific ways. That claim is resisted by a number of theorists who's experiments seem to show that emotions can be caused by a stimulus without an intervening process of stimulus evaluation. For example, Zajonc (1980) conducted a series of experiments which showed that subjects can form preferences for stimuli to which they have been subliminally exposed. Subsequent confirmation of that result came in a series of experiments conducted by Öham and his collaborators (1993). In one of these experiments, participants were conditioned to dislike angry faces. Participants were then shown an angry face followed in quick succession by a neutral face. The conditioned emotional response was elicited by participants when the angry face was masked by a neutral face even though verbal reports indicated that the subject did not consciously register the image of the angry face. This research suggests that an emotional response to a stimulus can be elicited without conscious awareness of that stimulus. In a later study, subjects were exposed to subliminal images of snakes, spiders, flowers and mushrooms. Again, subjects were unable to identify which stimulus they had been exposed to when questioned, however it was shown that subjects with previously established snake phobias showed elevated skin conductance responses to masked images of snakes and that subjects with previously established spider phobias showed elevated skin conductance responses to masked images of spiders.²⁴ As previous evidence has shown that elevated skin conductance responses are a marker for emotional arousal, these

²² In sadness evaluations of future expectations are left open.

²³ Lazarus 2001 p. 64

²⁴ Öham & Soares 1994; Öham 2002

results confirm that a physiological marker of a state of emotional arousal can occur without the subject being consciously aware of any evaluations of the emotion's stimulus.

The examples so far suggest that it is not necessary for a subject to be aware of an appraisal of a stimulus in order to emotionally react to it, or more exactly, in order to have one of the physiological markers of emotional arousal. There are a range of other examples supporting this conclusion. Appraisals, it would seem, are neither necessary nor sufficient for emotional arousal. Appraisals are not sufficient because a subject can change her explicit appraisals without changing her emotional response. For example, a perception of a spider may elicit a fear response in a phobic patient even if she does not appraise that object as dangerous (as when she does not consciously register the presence of the spider) or she appraises her situation as one which does not instantiate danger (as when she consciously registers the presence of the spider but does not appraise it to be dangerous). Appraisals are not necessary because emotions can occur independently of and prior to cognitive appraisals. An emotional response will occur independently of a cognitive appraisal when it is caused directly by physical means (e.g., by ingesting a drug). An emotional response will occur prior to a cognitive appraisal when that response occurs so quickly that a subject does not have the time to perform an evaluation of it. Consider for example, a fear response experienced by a rabbit confronted with a snake. Zajonc (1980 p. 156) writes that, "if the rabbit is to escape... the action must be undertaken long before the completion of even a simple cognitive process- before, in fact, the rabbit has fully established and verified that a nearby movement reveals a snake in all its coiled glory." So, it seems *either* that the fear response and its associated flight behaviour occurs before an evaluation of the stimulus *or* that an evaluation of the stimulus occurs before perceptual processing of that stimulus has been completed. Either conclusion seems untenable. Surely, the mental processes essential to an elicitation of an emotional response cannot be thought of as an evaluation of a stimulus if they can generate that emotional response *before* a subject has perceptually processed it. That is, before the perceptual information derived from the stimulus makes its way into consciousness.

It may have been noticed that the counterexamples presented are convincing precisely because in the cases described it is unclear that a subject could *consciously* and *deliberately* make the evaluations of the stimulus supposedly required to experience them. For example, it seems a little strange to say a subject consciously evaluates a masked image as dangerous when she is not consciously aware of that image and it is a little strange to suppose that a rabbit deliberately evaluates a snake as dangerous when, in order to survive, the rabbit's fear response must be activated before perceptual processing of the snake is completed. Equally, it seems strange to insist that a phobic patient who believes

that this spider is benign must consciously evaluate it as dangerous. The task for an appraisal theorist is clearly to accommodate the evidence which suggests that *awareness* of an evaluation of a stimulus is not necessary or sufficient to react emotionally to it. To do this, many appraisal theorists simply deny that appraisals, qua cognitive processes, must be slow, deliberate and conscious. Lazarus explicitly states that it is not necessary that a subject is aware of the appraisals that she makes and which (according to the hypothesis) are required to elicit an emotional response. Because appraisals may be automatic and unconscious, Lazarus claims that the experiments cited by Zajonc do not show that emotions can be elicited absent any cognitive appraisal at all. It *could* be that the appraisals involved in these cases are automatic and unconscious. Granting that appraisals may be automatic and unconscious, an appraisal theorist can accommodate all of Zajonc's examples only if the evaluations which generate an emotional response can be (i) capable of being completed without the information evaluated being available to other cognitive processes, as in the case of the phobic patient, (ii) capable of being completed before perceptual processing of the stimulus has been completed, as in the case of the rabbit and the snake, and (iii) capable of being completed using only simple sensory concepts to define the evaluative property that has been identified, as in the case of the snake and the masked images. But what's wrong with that? Zajonc writes that to understand evaluations in this way is to:

“[broaden] the definition of cognitive appraisals to include even the most primitive forms of sensory excitation, thus obliterating all distinction between cognition, sensation and perception.”²⁵

This description of the problem does not make explicit what is wrong with Lazarus' response. The term “appraisal” is a technical term and as such it may be defined by a theorist in whatever way he chooses. The problem with Lazarus' response is not that he expands the notion of an appraisal to include (i)-(iii) but rather that as a result of his definition, the appraisals leading to an emotional response are to be thought of as two independent processes. There are those appraisal processes that run along neural pathways which generate quick, automatic and unconscious responses to a stimulus and those that run along neural pathways that generate conscious, deliberate and slow assessments of a stimulus. These two processes are independent of each other: they may occur alongside each other and they may contradict each other so that the “actual” automatic appraisals which generate an emotion may not only fail to register with a subject they may well run contrary to the subject's explicit appraisal of the stimulus. To use the term appraisal to designate both sorts of processes is unhelpful because it runs the

²⁵ Zajonc 1984 p. 117

risk of equivocation. Take for example, the primary means of distinguishing between self-conscious and basic emotions. It is no coincidence that none of the experiments which Zajonc cites against AT are about self-conscious emotions: these emotions tend to be generated independently of conscious and deliberate forms of information processing and so it is highly likely that *if* self-conscious emotions are caused by appraisals, at least some of those appraisals will be conscious and deliberate. Indeed most appraisal theorists talk as if the appraisals involved in the elicitation of self-conscious emotions are invariably conscious and deliberate. Recall Tracy and Robins (2004 p. 105) claim that:

“Although basic emotions such as fear and sadness can and often do involve self-evaluative processes, only self-conscious emotions must involve these processes.... A sense of self, as conceived by self theorists since James (1890), includes an ongoing sense of self-awareness (the "I" self) and the capacity for complex self-representations (the "me" self, or the mental representations that constitute one's identity). Together, these self-processes make it possible for self-evaluations, and therefore self-conscious emotions, to occur.”

The claim here is simply that the sorts of self-evaluations which enable self-conscious emotions to occur require a subject to direct attention onto a self representation. The critical point is that directing attention onto a self representation is an ability that involves complex, deliberate and conscious forms of information processing.²⁶ As making a self-evaluation involves assessing a self-representation in light of evaluative criteria, it is very unlikely that making a self-evaluation is a process that can happen offline. So it seems that the way in which basic and self-conscious emotions are distinguished here supports the view that the evaluations involved in self-conscious emotions must be deliberate and conscious whereas the evaluations involved in basic emotions need not be. If self-conscious and basic emotions are distinguished modally by the sorts of appraisals they involve, the claims that both self-conscious and basic emotions require appraisals for their elicitation is likely to mislead. When we keep the processes designated by the term “appraisal” constant, it cannot be true that self-conscious emotions require appraisals for their elicitation and that basic emotions require appraisals for their elicitation. It is to avoid the danger of equivocation that Lazarus ought to have defined his terms less broadly.

²⁶ Tracy & Robins 2004; Lewis 2000; Izard *et al.* 1999

3. Empirical Support for Appraisal Theory?

In spite of the dangers of equivocation, proponents of AT are obviously aware that different sorts of evidence would be required to support the claim that unconscious appraisals cause emotional arousal and the claim that conscious appraisals cause emotional arousal. In this section, the empirical evidence invoked in support of AT is discussed. I raise some methodological problems associated with the experiments used to gather evidence for the specific correlations between emotions and appraisals posited by some appraisal theorists. I conclude that these worries weaken the import of any empirical evidence invoked to support the claim that shame is caused by self-evaluations. Further, I argue that whatever empirical evidence is invoked to support the claim that shame is caused by self-evaluations cannot show that shame is not also caused, perhaps quite frequently, by subjects who do not make a self-evaluation. This final point will set us up for a discussion in chapter 2 of cases in which shame is experienced by a subject who makes no self-evaluation at all.

The history of the attempt to empirically support the idea that subjective evaluations intervene between a stimulus and an emotional response goes something like this. First, Speisman *et al* showed that different interpretations of the same visual stimulus can influence the stress experienced as a response to that stimulus.²⁷ Here the stress response was not measured by verbal reports but rather by physiological indicators of stress such as skin conductance responses. In one of the experiments participants were asked to watch a film showing what looked like a painful operation on the genitals of a young man in an African tribe. The purpose of the experiment was to determine whether by manipulating the commentary which accompanied a distressful film the experimenters could modify the stress experienced by participants. Using no commentary as the measure of a control condition, they found that when no commentary accompanied the film or when the associated commentary emphasised the pain of the subjects involved (trauma commentary) participants experienced significantly higher skin conductance responses than when the commentary suggested that the ritual was painless (denial commentary) or when it encouraged the participants to take a detached view of the event by thinking of it from a scientific or anthropological perspective (intellectual commentary). These results show that the associated commentary could be used to determine the level of stress experienced in response to the film. It wasn't a huge leap from this result to the hypothesis that the commentary altered participants stress levels by altering their interpretation of the film. Thus it was concluded that the objective characteristics of a stimulus were not as important to determining emotional arousal as the *evaluations* (i.e., interpretations) of

²⁷ Speisman, Lazarus, Mordkoff, & Davison, 1964

those objective characteristics.²⁸ This conclusion was confirmed by Lazarus in a series of experiments which added to the research conducted by Speisman *et al* by varying the moment at which the commentary was given.²⁹ Lazarus found that denial information had the greatest affect on lowering stress when it was given *in its entirety before* participants watched a distressful film. This prompted the suggestion that when information relevant to a subjects interpretation of the meaning of the film is given beforehand it enables a subject to better manage future expectations by forming beliefs about what she will see and about how those images should be interpreted. The picture which emerges from both Lazarus and Speisman *et al* was that beliefs about the film (which are manipulated by a commentary) help to determine a subjects' interpretation of it and a subjects' interpretation of the film helps to determine levels of emotional arousal experienced in response to it. This provides empirical support for the idea that an emotional reaction to a stimulus is influenced by subjects' beliefs about that stimulus. Of course, it does not provide support for the claim that specific evaluations are connected with specific emotions.

The evidence for a direct correspondence between specific evaluations and specific emotions originally came from an attempt to confirm or disconfirm model's of emotional causation. What happened was that due to the increasing evidence of the effect of cognitions on emotions, model's were proposed which identified a number of appraisal dimensions and which made a number of predictions as to how evaluations along these dimensions of appraisal led to specific emotion types or how differences in the specific evaluations made along these dimensions can be used to distinguish one emotion type from another. Experiments were then constructed which either confirm or disconfirm these predictions.³⁰ For example, Smith and Ellsworth (1985) conducted an experiment in which participants were asked to recall one event which led to an emotion *x* for fifteen different values of the variable *x*. Participants were then asked to respond to a series of questions designed to provide a measure of the relevance of eight appraisal dimensions on the elicitation of each specific emotion. They found that the appraisals identified by subjects for each of these fifteen emotions closely corresponded to 6 of the hypothesised appraisal dimensions suggested by their model. They also found that the model successfully predicted which set of evaluations were crucial to the elicitation of some

²⁸ The option that the commentary itself constituted an objective characteristic of the film should here have been considered.

²⁹ Lazarus & Alfert 1964

³⁰ Many of the hypothesised models differ in the view they take about the relevance of specific appraisal dimensions to the production and differentiation of particular emotion types. These differences will not concern us here because any empirical data invoked to substantiate a specific model supports the more general claim that an identifiable set of evaluations generates a particular type of emotional response. It is support for the general claim that is of interest.

among these emotions. For example, evaluations of responsibility and control were correctly predicted to be crucial to the elicitation of anger whereas evaluations of self-responsibility and control were correctly identified as crucial to the elicitation of shame and guilt. In a later study Scherer and Ceschi (1997) showed how different appraisals can generate different emotional reactions to the same event. Here, the researchers tracked the emotion experiences and the antecedent evaluations made by a group of travellers a short time after losing their luggage at a major international airport. The method of data acquisition here was a structured interview designed to elicit information about the participant's appraisal of the situation as well as their actual emotional reactions to the news that their luggage had been lost. Results showed that appraisals of the loss as highly obstructive of goals and as incompatible with norms typically produced anger, although these appraisals could not be used to predict the intensity of the anger experienced.

Both Smith and Ellsworth (1985) and Scherer and Ceschi (1997) conduct experiments which appear to support the existence of a correlation between appraisals and specific emotions. I will label the method used by these theorists to acquire evidence for the correlation between appraisals and specific emotions *the subjective recall method*. The subjective recall method requires participants to reflect on an *actual* emotional event after it has occurred and then to answer questions about the evaluations which, as they see it, led to that emotional reaction. One of the worries with using the recall method to establish claims about a correlation between specific emotions and appraisals is that a participant's response may be due to a post hoc interpretation or rationalisation of the recalled experience. In fact, such rationalisations are actually encouraged by the sorts of questions that subjects are required to answer in these experiments. For example, a subject who described a circumstance in which she experienced anger may be asked, "do you think that you are to blame for that circumstance?" and "were you offended by that circumstance?" These questions already point in the direction of the appraisal dimensions that would rationalise the subjects' emotional response. It is not only the line of questioning which prompts a subject to a post hoc rationalisation of her experience but also the fact that, in general, we tend to understand our emotional behaviour in a way that fits into a positive narrative of our lives. This again supports the view that, our answers to those questions which are designed to bring into view the causal antecedents of recalled emotions will be skewed by retrospective interpretations of that emotional response and the evaluations leading up to it.

Acknowledging the shortcomings of the subjective recall method, Smith and Lazarus (1993) used *the vignette method* to gather support for a correlation between certain appraisals and certain emotions. This is a method in which participants are asked to read a

series of short stories about a character involved in some event and to answer questions about what the character to whom the event occurred would feel. For example, participants may be presented with a story in which a subject has preformed badly in an important exam and she attributes her failure entirely to the inadequate teaching that she received (thereby implicitly stressing an appraisal of other accountability). Participants are then asked whether this character would feel emotions *x, y, or z* in that situation. Smith and Lazarus use the vignette method to show that by manipulating appraisals they could affect the emotions which their participants predicted would be experienced by the subject in the story. In their particular experiment, participants were presented with scenarios and asked to adopt the perspective of the central character in the scenario. From that perspective, participants were required to answer a series of questionnaires, one of which was to determine the emotion experienced by the subject in the scenario given the situation that was therein represented. The results showed that the manipulation of appraisals did have an effect on the emotional states reported by subjects to be experienced by the character and that the model proposed by Smith and Lazarus which had 6 appraisal dimensions, for the most part, successfully predicted the effect that the manipulation of appraisals would have on the emotional states reported. For example, it was successfully predicted that anger was more common when there was other-accountability rather than self-accountability and that guilt was more common when there was self-accountability rather than other-accountability.

Here again we might raise a methodological objection. The worry now is that we cannot substantiate any causal claims on the basis of the evidence obtained by either the vignette method or the subjective recall method. Why? Results obtained from the subjective recall method only warrant the conclusion that certain past emotions are believed to be correlated with certain appraisals and as such they are insufficient to establish that emotions are caused by evaluations. Similarly insufficient are the results obtained from the vignette method which only warrant the conclusion that a person who imagines a subject who makes certain evaluations about an identified event will typically believe that the imagined subject experiences certain emotions. It seems then that neither of these two methods provides substantial evidence for AT which is at heart a theory constituted by the causal claim that emotions are *caused* by subjective evaluations.³¹ In an attempt to show that appraisals and emotions are causally related and that the direction of causality is principally from appraisal to emotion, Roseman and Evdokas (2004) propose a different methodology.

³¹ Roseman & Evdokas 2004 p. 1

What we might call *the elicitation method* monitors the effect that manipulating appraisals has on actual emotional experiences and as such it is better able to lend support to the claim that emotions are caused by appraisals. Using this methodology Roseman and Evdokas looked at the effect that manipulating two appraisal dimensions would have on an experience of three positive emotions: joy, relief and hope. The two appraisal dimensions that were manipulated in their study are *motivational state* (i.e., an evaluation of whether a subject wants more or less of what's at stake) and *probability* (i.e., a perception of whether what is at stake is either uncertain or certain). Participants in the experiment were split into two groups: a pleasant taste group were told that they would receive a pleasant taste or no taste at all and an unpleasant taste group were told that they would experience an unpleasant taste or no taste at all. Within the pleasant taste group some were told that they would definitely be in this group and others were told that they would probably be in this group. Similarly, within the unpleasant taste group some were told that they would definitely be in this group and others were told they would probably be in this group. We are to suppose that information about the group to which one is assigned provides a measure of motivational state because ordinarily that information will invoke a desire either to obtain a pleasant taste or to avoid an unpleasant taste. We are also to suppose that information about whether ones membership to a particular group was definite or merely probable provides a measure of probability because this information will ordinarily invoke evaluations of certainty or uncertainty (e.g. I will certainly be receiving an unpleasant taste). The results of the experiment showed that, in line with predictions, (1) participants who were led to believe that they would experience a pleasant taste rather than an unpleasant taste reported relatively high levels of joy, (2) participants who were led to believe that they had definitely avoided an unpleasant taste reported relatively high levels of relief and (3) participants who were led to believe that a pleasant taste would probably be experienced reported relatively high levels of hope. These results are supposed by researchers to provide evidence for the claim that appraisals are among the causal determinates of an emotional experience. Are they right? Have we arrived at some decent empirical evidence for AT?

4. Methodological Worries

In all of the three methods considered above, conclusions about the relationship between emotions and appraisals are reached on evidence obtained from subjective reports. More specifically all of the evidence for AT considered above was derived from participants who are asked to report on the evaluative processes that lead to an actual or imagined emotion

or to identify the actual or imagined emotion experienced by a subject on the basis of information given about the evaluations that that subject makes. Let's call any method which is supposed to provide empirical evidence for AT by using subjective reports *the subjective method*. Evidence for AT is obtained by the subjective method if it is acquired by asking subjects questions designed to determine their beliefs about the role that certain appraisals play in the elicitation of certain emotions. I take it that the subjective method is a method which encompasses other methods including the elicitation method, the vignette method and the subjective recall method.³² There are two problems associated with the use of the subjective method both of which suggest that results obtained from that method have little or no bearing on the claim that emotions are caused by evaluations. The first problem is that by relying on the subjective method there is the danger that appraisal theorists are doing nothing more than elucidating aspects of a folk psychological theory of emotions. If that is right, results obtained by the subjective method will be insufficient to establish the conclusions we are actually interested in. These conclusions have to do with the evaluative processes that really cause an emotion to unfold and not with the evaluative processes people *tend to believe* cause an emotion to unfold. The second problem comes from empirical studies which have already been considered which show that evaluations that lead to an emotional response can deviate from and even contradict the evaluations of that same stimulus that are verbally reported and integrated with a subject's conscious beliefs. Again if this is right, subjective reports may be an utterly unreliable source of evidence for supporting the claims that an appraisal theorist is interested in making, i.e.,

³² It may be argued that none of the three methods cited are actually attempts to ascertain a subject's beliefs about which set of appraisals cause which emotions. For example, with the elicitation method, the experimenters are trying to manipulate participants' appraisals of a situation and then asking them to identify which emotions they feel. This does not seem to be a way of ascertaining the participants' *beliefs* about which appraisals cause which emotions. Equally, in regards to the recall method, one might argue that the experimenters are trying to determine what emotion was felt and the sorts of evaluations with which that emotion was associated. Using this data, it is the *experimenters* who posit a causal connection between *this* specific set of appraisals and *that* particular emotion. If that is right, it is not, as I have suggested, a subject's beliefs about which appraisals cause which emotions that the experiments are designed to uncover but rather her *beliefs* about the appraisals she makes when experiencing a specific emotion. I cannot fully discuss these points here, however it should be noted that although the subjective method is supposed to cast doubt on the elicitation method, the recall method and the vignette method independent problems were raised for both the vignette method and the recall method. The worry with the vignette method and the recall method was that we cannot substantiate any causal claims on the basis of evidence obtained by those means. At best the use of these methods establishes a correlation between appraisals and certain emotions but there are various ways of explaining this correlation. If for example, we take the correlation to indicate a causal relation why not take it that the direction of causality is from emotion to appraisal? These problems would suggest that if we are to take seriously the causal claims made by an appraisal theorist, our best methodology would be the elicitation method. Now, if we used the elicitation method to show that the subjects who are encouraged to evaluate themselves negatively in a given situation experience shame and subjects who were not encouraged to evaluate themselves negatively did not experience shame, the story could not end there. Such an experiment would not show that shame is not also experienced in cases in which no negative self-evaluation is made. Indeed the purpose of chapter 2 is to highlight just that sort of case. In other words, although we may well use the elicitation method to provide empirical evidence for the claim that shame is caused by self-evaluations, this evidence *could not* be used to support the claim that self-evaluations are *required* to experience shame because there are cases of shame which do not involve negative self-evaluations; cases which it is the purpose of chapter 2 to discuss.

claims about the actual relations between evaluations and emotions. In what follows I focus on the first worry, as I take it that bringing the second worry into view only requires us to recap the sort of experiments conducted by Zajonc and discussed in the previous section.

To begin with a sketch of the first objection, let's get in view what folk psychology actually refers to in this context. Folk psychology is a psychological theory of the mind that is constituted by the platitudes that ordinary people are likely to endorse about the mind and its operations. More specifically, following Lewis (1972) we can think of folk psychology as a collection of

“[A]ll the platitudes you can think of regarding the causal relations of mental states, sensory stimuli, and motor responses. Perhaps we can think of them as having the form: *When someone is in so-and-so combination of mental states and receives sensory stimuli of so-and-so kind, he tends with so-and-so probability to be caused thereby to go into so-and-so mental states and produce so-and-so motor responses. Also add all the platitudes to the effect that one mental state falls under another—“toothache is a kind of pain” and the like.* Perhaps there are platitudes of other forms as well. Include only platitudes which are common knowledge among us—everyone knows them, everyone knows that everyone knows them, and so on.” (Lewis, 1972 p. 207–8. Italics added)

The worry is that the results accumulated by the appraisal theorist on the basis of verbal reports seem rather like the beliefs that are constitutive of folk psychology. Isn't it just a matter of common sense psychology that joy is more likely when a subject is led to believe that she will eat something pleasant rather than when she is led to believe that she will eat something unpleasant (Roseman and Evdokas 2004), or that anger is more likely when there is other accountability rather than self accountability (Smith and Lazarus 1993), or that shame and guilt are more likely when there is self accountability rather than other accountability (Ellsworth and Smith 1985)? The problem is that many of the beliefs of which folk psychology is comprised are beliefs about which appraisals are appropriate to which emotions and which situations will elicit certain appraisals. The danger is as Griffith puts it: “failure to distinguish between elucidating the folk theory and studying emotion processes themselves is unlikely to lead to a good account of either.”³³ But when we distinguish folk psychological claims from facts about emotions, it becomes apparent that evidence received from the subjective method is simply insufficient to establish any

³³ Griffith, 2003 p. 47

conclusions about what goes on in the actual evaluative processes leading to an emotional response. Thus it seems that:

- (1) When the evidence for a causal relationship between a set of appraisals and an emotion is obtained by the subjective method that evidence succeeds in doing nothing more than reflecting commonly held beliefs which subjects have about emotional elicitation.

An elucidation of a folk psychological theory of emotions may be relevant to gaining a better understanding of the emotions but it is an elucidation of the actual evaluative processes that lead to an emotional response that an appraisal theorist is interested in. That is to say, if the evidence obtained by subjective reports are to be taken seriously as evidence for AT rather than merely as reflections of some aspects of our folk psychological understanding of emotional causation, some additional work is required. For example, we will need good reason to think that the results obtained by the subjective method (which as we have seen pertains to the specific appraisals that are *commonly believed* to be associated with specific emotions and the circumstances in which those appraisals are appropriately made) *for the most part* reflect facts about the evaluative processes that lead to an emotional response. That is to say, if we are to take as evidence for AT whatever relationship between appraisals and emotions that is reflected in commonly held beliefs about emotional causation, a discernable and positive relation must be established between those beliefs and the facts they purport to represent. But that is not all. The fact that we have this empirical method at all presupposes that the thing about which we are acquiring evidence (i.e., AT) is something for which we *could* acquire *empirical* evidence. Thus the evidence about emotion elicitation that is acquired by the subjective method had better be empirical evidence- which is to say, it had better be comprised of facts that cannot be arrived at simply from an understanding of the meaning of the relevant emotion terms.³⁴ In fact, this requirement is far more important than establishing a point of contact between the results obtained by the subjective method and facts about the processes that lead to an emotional experience because if the evidence acquired by the subjective method is comprised of facts which can be known on the basis of words alone then it is a thankless task constructing experiments to accumulate such facts. More specifically, investigation into the relationship between appraisals and emotions by empirical means would literally be a waste of scientific resources as a theorist with enough imaginative dexterity could, sitting at a desk, think up a range of situations and use her knowledge of the meaning of emotion terms to determine which appraisals of that situation will lead to which emotions.

³⁴ An emotion term is a word used to designate emotional experience (e.g. *joy, anger* and *fear*)

We can put the points made so far as follows. In realising that the evidence obtained by a subjective method merely reflects commonly held beliefs about the sorts of evaluations that cause particular emotion types such as “anger”, “shame”, “fear” and the like, the most pertinent question is *not* what is the relationship between these beliefs and facts about the cognitions which trigger the onset of an emotional process but rather *are the claims about the relationship between appraisals and emotions which are reflected in these beliefs claims that can be arrived at simply by thinking about the meaning of words?* The reason for asking this question first is that the answer to it determines whether a study of the relationship between appraisals and emotions using the subjective method is a worthwhile *scientific* project. Unfortunately for an appraisal theorist things do not look very promising when we look again at the sorts of results obtained by the subjective report method. For example, the idea that joy is more likely when one expects to experience a pleasant taste rather than when one expects to experience an unpleasant taste is a fact which follows from knowing the meaning of “joy”. We simply would not know what joy meant if we thought that the prospect of receiving anything unpleasant, without any further information, could be a cause of it. Nor would we know what *relief* meant if we did not expect it to be experienced by a subject who is led to believe that she has definitely avoided the prospect of tasting something unpleasant. It is a fact about the meaning of “relief” that it is a sort of state experienced when something does not happen which could have happened and which one does not want to happen. Similarly one could not know what anger meant if one did not know something of the appraisals which typically elicit it. For example, if I did not know that anger is an appropriate response to a variety of situations for which I am not responsible (like the failure of a gas man to come at an agreed time) I just wouldn’t know what anger meant as I would not have sufficient knowledge about the sorts of circumstances in which it is appropriate to apply that word. The same goes for the results obtained about shame and guilt. One could not understand what “shame” or “guilt” meant if one did not know that it is typically experienced when one blames oneself for some wrong doing. It seems then that the results considered here, which have been empirically confirmed by AT using the subjective report method, could have been secured simply by knowing the meanings of the relevant emotional terms. That is to say:

- (2) The sort of conclusions reached by the subjective method not only appear to reflect commonly accepted platitudes about emotional elicitation, they seem to reflect platitudes which can be known on the basis of an understanding of the meaning of emotion terms.

Let's sum up the worry briefly. Any empirical method that is used to test the viability of AT will itself make reference or require its participants to make reference to emotional phenomena such as "joy", "shame", "guilt" and so on. But in making reference to just these terms experimenters presuppose that participants have some adequate grasp of the ordinary meanings that these terms have. I assume that knowing the ordinary meaning of an emotional term involves knowing something about the circumstances in which it is appropriate to use them. This knowledge in turn requires knowing something of the appraisals which would warrant the use of that term in just these circumstances. But if having an adequate grasp of the meaning of emotional terms entails knowing something about the sorts of appraisals which make the application of those terms appropriate in some given context, then using an empirical method to investigate the relationship between appraisals and emotions is superfluous. These relations may just as well have been established by an investigation into the meaning of emotional terms. From these observations we may conclude either that:

- (3) When relying on a subjective report method, AT does "little more than explicate the implicational semantic structures of our emotion vocabulary"³⁵ or that AT is "misleadingly or incorrectly empirical in that it seeks empirically to test what follows from the meaning of words"³⁶

We can call (1) – (3) *the no-relevance argument* because it questions the relevance of any investigation which seeks to confirm AT solely on the basis of the subjective report method. Before recapping this argument, I wish briefly to set aside an objection to it. The objection is that if the experiments conducted by an appraisal theorist merely reflect facts about the meaning of terms, how is it that competent speakers' can disagree with each other about the set of appraisals required to experience certain emotions? In fact, aren't all the disputes amongst appraisal theorists primarily disputes about which appraisals are required to elicit specific emotions, and aren't these appraisal theorists competent users of language? This sort of observation requires us to be a little more careful in stating the worry that the claims made by an appraisal theorist as a result of his various experiments are analytic. There are a set of appraisals which are connected with emotions which seem to be based on word meaning alone. To determine which sets of appraisals are connected to emotions through meaning we may ask, can a competent speaker of the language use this term in such a way as not to presuppose that *this particular* appraisal is being made? To illustrate that point, take the example of anger. Two subjects may differ on what they take a perceived offense to be, however they cannot disagree on the fact that "anger" is

³⁵ Scherer 1999 p. 655; Parkinson 1995 p. 47-8

³⁶ McEachrane 2009 p. 34

correctly applied when there is a perceived offence. As a matter of meaning, anger is connected to an appraisal of a situation as involving a perceived offense. The point is that, in all of the examples that we have seen, the sorts of appraisals thought to have been uncovered through experiments are of the sort that appear to be connected to emotions via their meanings. Competent speakers may well disagree on whether a situation can be classed as say, an offense, but they cannot disagree on whether a perceived offence is required to experience anger- at least, not without failing to understand the meaning of anger.

With this clarification in view, let's recap the no-relevance argument. In essence the argument goes like this: The evidence which is accumulated in support of AT derives from the subjective method which, on the face of it, does nothing but reflect platitudes that ordinary people are likely to endorse about emotional elicitors. From *that* evidence we can only come to an understanding of the appraisals that people commonly believe are associated with certain emotions. Looking at these commonly held beliefs we find that the claims we are licensed to conclude on their basis are claims about the relationship between specific appraisals and emotions which follow from knowing the meanings of the relevant emotional terms. The objector continues with the question: isn't an empirical experiment pointless if all it does is confirms facts that are based on the meanings of words? Assuming that the answer is "yes" the objector concludes: the very project of investigating AT by the subjective method is pointless.

In response to this objection, it is well worth asking the objector why we should agree that an empirical investigation into the relationship between appraisals and emotions would be pointless if all it did was elucidate facts that can be known by conceptual analysis. Surely, if we can make explicit aspects of the structure of particular emotional terms using the subjective method we will have done something useful. Furthermore going into the world to see how emotional language is actually used seems to be a great way of explicating the ordinary meanings of emotional terms. It is possible that such an empirical investigation into conceptual relations will lead to important discoveries because there is no reason to think that it will not bring to light surprising or overlooked aspects of the conceptual relationship between emotions and appraisals. In short, it is open to an appraisal theorist to respond to the no-relevance argument by conceding that in using only the subjective method to collect evidence about the relationship between appraisals and emotions he is in fact engaging in a conceptual investigation by empirical means. It does not follow from this that AT is pointless. Studying the relationship between appraisals and emotions on the basis of the subjective method may be justified simply because it renders our emotional concepts more precise. Interestingly, this way of

responding to the no relevance argument contradicts the popular view that when considered as a science, psychology is invariably an investigation into empirical rather than conceptual facts. It seems to me that this view is a prejudice. Of course conclusions about the structuring of concepts will have no immediate bearing on the reality of emotional processes but that does not mean that these conclusions are not important or that they should not be investigated by the psychologist. It only means that some psychological investigations are more akin to philosophical analysis than to empirical analysis.

A less unusual response to the no relevance argument would be to point out that, it cannot be the case that all of the connections between evaluations and emotions that are discovered to exist through experiment are the sorts of relationships between emotions and evaluations which can be known on the basis of the meaning of the relevant emotional terms. All we need to do is to apply our test. If a subject can be said to meaningfully use an emotion term in a situation in which the relevant appraisal is not present, then that appraisal is not presupposed in the meaning of the emotion term. If, therefore we conduct an experiment that established a connection between an emotion and *that particular* evaluation, we would indeed have some empirical evidence for the connection. Further, if when conducting that experiment the appraisal theorist uses the elicitation method, we would have some reason to think that the connection between the appraisal and the emotion is causal and that the direction of causation goes from appraisal to emotion. Unfortunately, the story cannot end here.

Consider the case of shame. According to an appraisal theorist an experience of shame is connected to some sort of negative self-evaluation. I take it that the word "shame" can be sensibly applied in situations in which a subject does not make a self-evaluation and as such it is possible to gain empirical support for an appraisal theorist's claim about shame. If the elicitation method is used to show that subjects' who are encouraged to evaluate themselves negatively in a given situation experience shame and subjects who were not encouraged to evaluate themselves negatively do not experience shame, we would have established empirical support for a causal connection between shame and a negative self-evaluation. The problem for an appraisal theorist who conducts such an experiment however, is that the cases he sets up in the experiment could not exhaust all of the situations in which a subject experiences shame. For this reason, an experiment which provides empirical support for the connection between shame and a negative self-evaluation *could not* show that shame is not also experienced in cases in which no negative self-evaluation is made. Indeed the purpose of chapter 2 is to highlight just that sort of case.

5. Chapter Summary

To summarise: AT is a psychological account of emotional elicitation which attempts to test by empirical means the claim that emotions are caused by subjective evaluations of a stimulus in light of an agent's own concerns. Two observations were made about this view. First, because there are two independent evaluative processes prompted by the emotion stimulus (one conscious and automatic, the other deliberate and unconscious) there is the danger of equivocation. For example, the claim that fear is caused by an appraisal and the claim that shame is caused by an appraisal may well mean different things depending on the sort of processes designated by the critical term "appraisal." If fear is caused by an appraisal that is automatic and unconscious, it is caused by a process of a kind that is quite different to the one involved in the elicitation of shame because the subjective evaluations involved in self-conscious emotions are, by all accounts, conscious and deliberate. Keeping in mind the danger of equivocation is important to stop an appraisal theorist and his opponent talking past each other. But that is not all. The two sorts of processes that the term "appraisal" may designate point toward a problem with using the subjective report method to investigate the evaluations involved in basic emotions. This is because, unlike self-conscious emotions, basic emotions such as fear, sadness and joy are more likely to be caused by appraisals that are unconscious and automatic. As we have seen, the unconscious and automatic appraisals which cause an emotional response may contradict the automatic and unconscious appraisals which lead to it. Because the appraisals which are verbally reportable are those which can be made conscious, the subjective report method can only point us in the direction of consciously available evaluations. This method is therefore liable to mislead when investigating the elicitors of basic emotions. In these cases the conscious appraisals arrived at may not be correlated with the evaluations that are actually responsible for their elicitation.

The second critical point made in this chapter concerned the methodology used by an appraisal theorist to support his claims about the causal relationships between particular emotions and specific appraisals. It was claimed that of the three methods discussed, only the elicitation method could provide evidence for a causal connection between appraisals and emotions. Furthermore, it was suggested that the elicitation method can empirically support claims about a causal relationship between specific emotions and specific appraisals only when those appraisals are not presupposed in the very meaning of the relevant emotional terms. So for example, if knowing the meaning of fear requires you to know that it involves an appraisal of danger, then the relationship between an appraisal of

danger and fear could not be empirically tested. Not unless, of course, it was admitted that the appraisal theorist was explicating a conceptual relationship by empirical means. To check whether an evaluation is presupposed in the meaning of a term, we had to ask whether or not that emotion term could be used by a competent speaker of the language in a way which did not presuppose that this particular evaluation is being made. For example, can anger be meaningfully used in a situation which did not involve an offence? If the answer is yes, then the relationship between an appraisal of offence and anger could be empirically investigated without the relevant admission. It was claimed that a negative self-evaluation is not presupposed in the very meaning of the word shame, and as such it is possible to use the elicitation method to acquire empirical evidence in support of the idea that shame is causally related to a negative self-evaluation. This empirical evidence however could not show that shame is not, and perhaps with great frequency, caused by situations in which no negative self-evaluation is made. In the next chapter, I will discuss a range of cases in which shame is experienced but where no negative self-evaluation is required *or* where the role played by a negative self-evaluation is not the one envisaged by most appraisal theorists.

Chapter 2

1. Weak and Strong Versions of The Dominant View

According to The Dominant View (DV), shame requires a subject who makes a negative self-evaluation. We can take it that a subject makes a negative self-evaluation either when she judges that she has failed according to some evaluative criteria or when she judges that she is x , where x , is a negative predicate such as bad, stupid, a failure and so on. The latter sort of judgment has been labelled a *global and punitive judgment*.³⁷ These judgments are global because they are directed to the self as a whole rather than a specific aspect of the self or the self's conduct and they are punitive because they are negative. Let's note however, that the choice of the word "punitive" points beyond the fact that these judgments are negative. The view that shame requires a self-evaluation which is global and punitive suggests to me a view according to which shame is a sort of penance, a self-inflicted punishment for a perceived wrong doing. This is a view of shame to which we will return. The focal point, for now is that because the negative self-evaluation thought to be required to experience shame can be understood in at least two ways, we have at least two different versions of DV in play:

The strong version of DV: the experience of shame requires a subject who judges that she *is* a failure (i.e., a subject who makes a self-evaluation that is global and punitive)

The weak version of DV: the experience of shame requires a subject who judges that she has failed according to evaluative criteria.

The second of these views is weaker than the first because a subject can judge that she has failed according to criteria without endorsing the judgment that she is a failure but she cannot endorse the judgment that she is a failure without recognizing that she has failed according to some criteria. In this section, I wish to discuss why it is that some proponents of DV have not been satisfied with its weak variant. This discussion has two points to it. First, it is the strong version of DV that has been championed by the appraisal theorist, and so the discussion will help us to better understand their motivations for preferring it. Second, the discussion will bring to light the possibility of a discord between the criteria which the subject employs to negatively evaluate herself and the criteria which the subject

³⁷ Lazarus 1991 p. 241

endorses. That discord will then be used in the following section to show that the strong version of DV is mistaken.

I want to discuss two sorts of reasons for thinking that the weak version of DV is inadequate. The first is that the phenomenology of shame demands that the self-evaluations it involves are global and punitive. Williams (1993 p. 89) aptly describes the phenomenology of shame when he writes:

“[t]he expression of shame, in general, as well as in the particular form of it that is embarrassment, is not just the desire to hide, or to hide my face, *but the desire to disappear, not to be there.* It is not even the wish, as people say, to sink through the floor, but rather the wish *that the space occupied by me should be instantaneously empty.*” (Italics added)

If a feeling as intense as this, a feeling which expresses itself partly as a desire for self extinction, requires a negative self-evaluation at all, surely it would be an evaluation of the type that one is unworthy, small, bad and so on. The evaluation that I have failed according to criteria simply doesn't capture or warrant or cohere with the extremity of feeling that shame is associated with. Helen Lewis (1971 p. 30) makes a similar suggestion. According to Lewis, the sorts of evaluations involved in shame are global and punitive. Lewis claims that shame is an acutely painful emotion precisely because it involves evaluations of just this kind. In shame the self is under attack: one feels oneself exposed to be *worthless, a failure*, lacking in power and so on. Lewis argues that this feeling is rooted in a deeper conflict in the psyche than the feeling of wrong doing which is central to guilt. Guilt unlike shame does not involve a judgment about ones identity but only about ones behaviour. Because *behaviour* is the focus of attention in guilt, we can escape guilt by changing our behaviours but because the self is the focus of attention in shame we can escape shame only through self-annihilation: by ceasing to be the sort of person who performs such acts or who is shamed by such acts. This is precisely why guilt is less painful than shame, Lewis argues. Guilt involves the recognition, all be it painful, that one is responsible for a wrongdoing which warrants punishment. The wrong act, however, we take as separable from the self and so as redeemable: *“even the misery of guilt doth attain to the bliss of a pardon.”* Shame on the other hand is an attack on the entire self and as such all hopes of reparation lies in the extinction of that self. The judgment that I have failed according to some criteria simply does not have the gravitas to sit alongside this description of the phenomenology of shame. Surely, if an attack on the self is required to experience shame, that attack must be reflected in the sorts of negative evaluations that a subject endorses

when she is experiencing that emotion. The weak version of DV will not do then, because it does not explain or otherwise accord with the grave feeling of inadequacy that is associated with an experience of shame.

The worry with the argument just sketched is that it neglects the fact that the phenomenology of shame can be explained in multiple ways. So for example, shame is often associated with a fear of social exclusion. It is perfectly possible that fear of social exclusion is responsible for the anxiety felt when one experiences shame and that fear can just as well be appealed to explain why shame feels so bad. That is to say, there is no reason to suppose that the phenomenology of shame is ill served if the subject of shame merely judges that she has failed according to some evaluative criteria. This is because there are other features of shame which can explain its distinctive phenomenology and one such feature is the fear of social exclusion. The fear of social exclusion plays an indispensable role in theories of shame that emphasize its social nature. We can get at the social nature of shame by tracing the structure of that emotion to its manifestation in ancient shame cultures. These are cultures in which one's identity is constituted by one's membership to a group. The group itself is bound together by a set of values, norms and practices and membership to the group is evidenced by following the values, norms and practices which bind it. When one is seen to flout these group norms one becomes dishonoured. That means that one suffers a rejection from the group which amounts to an extinguishing of one's identity. Of course, ancient shame cultures do not fully characterise the structure of modern shame experience, however parallels can be noted. For example, typically the evaluative criterion against which one negatively evaluates oneself is derived from a social group. Because following the values of the group identifies me as one of its members, when I am seen to perform actions that are unfaithful to those values and I experience shame as a result, this may simply be because of a fear of the loss of social standing. The threat of rejection and hostility from the group which would otherwise be co-operative, can by itself explain the phenomenal aspects of shame.

There is, however, another sort of argument which suggests that the weak version of DV offers an inadequate characterisation of the evaluations involved in shame. The idea is that if we grant that the weak version of DV is correct, then there are other features of shame which, on pain of irrationality will force a subject to endorse an evaluation that is global and punitive. For example, most appraisal and attribution theorists argue that shame

involves an attribution of responsibility for the shame eliciting event.³⁸ A good representation of this view is Lewis, (2007 p. 133) who writes:

“I can, for example, evaluate *my* behaviour against *my* SRG [standards, rules and goals] and conclude that I have succeeded or failed. However, this will not lead me to either pride or shame unless I am prepared to believe that I am responsible for that success or failure.” (Italic’s added)³⁹

In other words, in order to experience shame I must not only judge that I have failed according to my standards, rules and goals, I must judge that I am responsible for that failure. It could be argued that when a subject takes responsibility for a shame eliciting event she will, on pain of irrationality, endorse a global punitive judgment. That is because it looks as though one could not have criteria according to which being an alcoholic is bad, judge that one is an alcoholic, and not think of oneself as bad. If that is the case, an appraisal of *self-blame*, understood as an attribution of responsibility for the shame eliciting event, bridges the gap between the judgment that I have failed to live up to my evaluative criteria and the judgment that I am *bad, unworthy, stupid* and so on, *presuming that the subject is rational*. The problem with this argument is pretty clear. For even if we assume that an appraisal of self-blame can be used to play a bridging role, an appraisal of self-blame is not required to experience shame. For example, I may be ashamed of my nose, my country or my poverty. These are things (or could be things) that I recognise are beyond my control and yet I may have a standard that I fail in virtue of having this bulbous nose, or in virtue of being so poor, or in virtue of being British. In all of these cases, although I recognise that I have failed some standard in virtue of having *x* or being *x*, I do not take myself to be responsible for having *x* or for being *x*.⁴⁰ Here then, the notion of self-blame cannot play a bridging role because the notion of self-blame is absent.

There are, however, other options. For example, Deonna and Teroni (2009) argue that a subjects’ identity, understood as her conception of herself, is partly constituted by the values she holds. These values shape the expectations she has with regard to herself and to others. They therefore inform the criteria on the basis of which she evaluates herself and others. So in shame, because the values which inform the criteria on the basis of which a subject evaluates herself are integral to her identity, a belief that she has failed to meet

³⁸ c.f. Tracy and Robins 2004

³⁹ Note: Lewis is an attribution theorist. The difference between AT of shame and an attribution theory of shame is that the former believes that the only appraisals relevant to generating shame (and self-conscious emotions more generally) are appraisals relating to the self.

⁴⁰ Contra Solomon 1976 p. 306; c.f. Weiner, 1985

those criteria will lead her to judge that she is a *failure, unworthy, bad* and so on. In other words, when the subject who experiences shame judges that she has failed according to criteria, she will also judge that she is a failure because of the connection between the values which underpin the criteria and her conception of herself. If this suggestion is right, we could say that a subject who in shame judges that she has failed according to criteria must, on pain of irrationality, accept a global punitive judgment. We can therefore bridge the gap between the weak and the strong versions of DV by adopting Deonna and Teroni's account. But there is a problem with their argument. It assumes that when a subject experiences shame, she must *endorse* the criteria according to which she is negatively evaluated. The subject is required to endorse these criteria because, on this account, she must accept the values which underpin them as integral to her identity. In the following section, I argue that this requirement is too strong as it would make recalcitrant shame impossible.

2. Recalcitrant Shame

We get at a definition of recalcitrant emotions, first by noting that there exists a normative relationship between an emotion and the way that a subject represents its intentional object. This normative relation is often expressed by saying that, in order for an emotion to count as say, fear, a subject must have specific beliefs about the significance of its object. As Hacker (2009 p. 13) puts it:

“[T]he agent must take the object of his emotion to satisfy the formal characteristics which determine the object as appropriate. If he fears A or A's action, he must believe that A and A's action are a threat. If he feels pity for another, he must believe that person to have suffered a misfortune. If he feels regret, remorse or guilt, he must believe that he has done something unfortunate or wrong.”

Without getting bogged down in a discussion of what sorts of evaluations are fitting or appropriate to which emotions, we can say simply that the object of fear is fitting if it is evaluated by the subject as *frightening*, that the object of anger is fitting if it is evaluated by the subject as *infuriating*, that the object of shame is fitting if it is evaluated by the subject as *shameful* and so on. With this in view, we can define recalcitrant emotions simply as follows: *an emotion is recalcitrant when a subjects' knows or believes that her evaluation of its object is not fitting.* For example, a subject who is afraid of spiders but who is aware

that spiders pose no danger to her wellbeing is experiencing recalcitrant fear. This is because, in spite of herself, the subject is afraid of an object which she judges is not *frightening*. Crucially, any emotion can be recalcitrant because any emotion can be experienced in the absence of a fitting evaluative judgment. Recalcitrant shame, for example, is just a case in which shame is experienced despite a judgement that the object of the emotion is not shameful. There are multiple real life examples of recalcitrant shame. In what follows, I will focus on the case of Tony. Tony is a young, gay man raised by Roman Catholics. He regularly attends church and while he doesn't agree with all of their teachings, he considers himself a devout catholic. Tony knows that there is nothing wrong with being gay, but he is all too often overcome with shame when he is attracted to a man. Tony's shame is recalcitrant because, in spite of himself, he is ashamed of an object which he knows is not *shameful*.

The existence of recalcitrant shame derails the argument of the previous section. To recap: the aim of the argument was to show that when we accept the weak version of DV we are rationally required to also accept the strong version of DV. The argument went like this: a subject must accept the criteria according to which she is negatively assessed when she experiences shame because of the relationship between the values which underpin that criteria and her conception of herself. It seems that if a subject accepts the criteria according to which she is negatively judged then she will be rationally required to accept a global punitive judgment. For example, if Tony accepts criteria which determine that being gay is bad, and he judges that he is gay, then it is rationally incumbent on Tony to judge that he is bad. In response to this argument, the case of recalcitrant shame was introduced. In recalcitrant shame a subject need not accept the criteria according to which a fitting evaluative judgment would be made. For example, a fitting evaluation for Tony to make when he experiences shame about his sexuality is that being gay is wrong. This is because, it is natural to suppose that shame directed at some state of affairs x ought to be related to a negative evaluation about x or (failing this) to an evaluation of x which does not contradict a negative evaluation of x . Tony, however explicitly denies that there is anything wrong with his sexuality and as such rejects the criteria on the basis of which a fitting evaluation would be made. If therefore, we insist (as proponents of DV do) that Tony is required to make a negative self-evaluation in light of his sexuality, it will not be the case that the criteria on the basis of which Tony makes this evaluation are ones he endorses.

The existence of recalcitrant shame undermines the argument of the previous section and it forces us to distinguish between two readings of the weak version of DV:

The weak version of DV (A): the experience of shame requires a subject who judges that she has failed according to evaluative criteria *that she does accept*

The weak version of DV (B): the experience of shame requires a subject who judges that she has failed according to evaluative criteria *that she does **not** accept*

As Tony does not accept the criteria on the basis of which being gay is wrong, and as it is his homosexuality that is the alleged source of his negative self-evaluation, the first reading of the weak version of DV is false. But that is not all. The existence of recalcitrant shame shows that the strong version of DV, which claims that an experience of shame requires a subject to make a global and punitive judgment, is false. Tony does not, in light of his sexuality, judge that he is a failure because he does not view his sexuality negatively. In fact, because Tony does not view his sexuality negatively, he is not required to make any sort of global and punitive judgment at all.

At this stage, one might insist that Tony *unconsciously* accepts a global and punitive judgment or that he *unconsciously* accepts criteria according to which being gay is wrong. This would mean that recalcitrant shame is somehow incoherent because the subject of that state would have to hold contradictory beliefs (e.g. a conscious belief that there is nothing wrong with being gay and an unconscious belief that being gay is wrong or a conscious belief that the spider is benign and an unconscious belief that the spider is dangerous).⁴¹ I take it, however, that imposing an unconscious belief onto the subject of recalcitrant shame is unacceptable. This is because, although recalcitrant emotions involve some sort of error, they are not reflective of an incoherent state of mind. The conflicts between emotions and judgments that recalcitrant emotions capture happen too often for it to border on the incoherent and in fact, they are readily intelligible: we can very easily understand Claire's fear of flying even though we know that she knows that flying is not dangerous.⁴² To have this understanding we do not need to suppose that Claire does not know her own mind (i.e., at an unconscious level Claire believes that flying is dangerous). It is equally unlikely that recalcitrant emotions can be sidelined as exceptional, inauthentic or *sui generis*. Unless we stipulate that emotions are genuine only if they have a fitting evaluative judgment, we have no reason to think that recalcitrant emotions are exceptional, inauthentic or *sui generis*. That is, unless we assume that being ashamed of *x* is by itself sufficient to count as judging that *x* is negative and reflects negatively on me, (a

⁴¹ Bartky 1990

⁴² Helm 2001 p. 42; Greenspan 1988 p. 18

point which would make it trivially true that recalcitrant shame is an exception to a general rule) we have no reason to think that recalcitrant shame is exceptional, inauthentic or *sui generis*. This is not to say that recalcitrant emotions are not puzzling. What is strange about them is not so much the absence of a fitting evaluative judgment but rather a subjects' recognition that her emotion lacks the relevant evaluative judgment. We think, if Tony knows that there is nothing wrong with being gay, why is he ashamed of it? If Claire knows that the spider is innocuous, why is she afraid of it? Questions like these add to the sense that there is in fact some normative principle governing the relationship between an emotion and a subject's beliefs about its object which cases of recalcitrant emotions fail to respect.

Most attempts to characterise the error involved in recalcitrant emotions claim that the subject of that state is simply, on some level, irrational.⁴³ This might seem obvious. It certainly seems irrational for Claire to be afraid of flying when she thinks that it is a perfectly safe form of transportation or for Tony to be ashamed of being gay when he does not think that there is anything wrong with it. What adds to the sense that these emotions are irrational is that when you ask subjects like Claire to explain the inconsistency of their judgments and their emotional reactions they say things like, "I can't help it, flying just terrifies me" or they affirm "it's just irrational"- which is to say, when questioned the subjects of recalcitrant shame give up any attempt at rationally justifying their emotional experiences. Could we not then say that subjects who experience recalcitrant shame are irrational and understand DV as providing a requirement for all instances of *rational* shame? Briefly, I want to suggest that we cannot. It is not true, I think, that all instances of recalcitrant shame are irrational. More specifically, I want to suggest a reason for thinking that when a subjects' shame is recalcitrant, and she judges herself negatively according to criteria she does not accept, her emotion will be rational if the criteria that she adopts are ones adopted by a representative member of her community. In these cases, recalcitrant shame is rational because it is an instance of the subjects appropriately responding to a failure to meet the standards and values of her community. To begin to make this case, we will need to look at the adaptive value of shame.

The consensus since Darwin (1872) is that emotional experiences have adaptive value. That is to say, emotional experience evolved to help animals deal with social and physical challenges that had implications for survival and reproductive fitness.⁴⁴ The basic idea is that the phenomenal, behavioural, motivational, physiological and expressive components

⁴³ Brady 2009

⁴⁴ Darwin 1872; Frijda 1994; Lazarus 1991; Ekman 1992; Gilbert 2007; Keltner & Buswell 1997

of an emotional experience helped animals to successfully navigate their physical and social environment which in turn enabled them to survive and to pass on these emotional capacities to their offspring. It is generally thought that each emotion has its own unique adaptive value. Fear, for example, evolved in order to motivate withdrawal from a situation that would lead to a physical threat and anger evolved in order to motivate attack when a situation involved a physical threat. Self-conscious emotions are of particular interest here because their adaptive value lies in their enabling animals to deal successfully with various social situations. As humans are social animals being able to maintain cordial relationships within the group would have been essential to survival, as only within the group could one form attachments (e.g. find a mate), engage in co-operative activities (e.g. hunt) and form other mutually beneficial inter-personal relationships (e.g. entertainment). Of course, to gain the benefits of group membership one must signal to others that one is a group member and as groups are marked by conventions and practices, a subject signals her membership to the group by enacting behaviours that are permissible according to those conventions and practices. To act in a way that contradicts the conventions and practices of the group is to advertise oneself as a non-conformer. It is to say, in effect, I am not part of *this* group. Non-conformity would have undermined survival either because it made one the target of collective aggression or because it led to ostracism from the group and therefore exclusion from co-operative ventures which were essential to survival. Hence Gilbert (2007 p. 285) writes:

“[B]eing rejected, shunned or expelled, or even being allocated an unwanted low social position in a social group, not only compromises reproductive chances... but activates stress systems and seriously compromises health regulating social relationships and survival”

Within this broadly functionalist framework, it is widely accepted that shame aided survival by drawing attention to the fact that a subject has done something which contradicted the practises and conventions that delineate group membership.⁴⁵ In effect, shame evolved to track those of our behaviours that would have been met with the threat of social exclusion. The sting of shame would have been adaptive because, for one thing, the adverse consequences which followed a transgression would serve as a lesson to conform to prevailing standards of behaviour, so that in time the anticipation of shame would prevent one from acting as a non-conformist.⁴⁶ Now, with the adaptive function of

⁴⁵ Tracy and Robins 2004; Leary 2007; Fessler 2004; Maibom 2010

⁴⁶ Fessler 2004 p. 241 Note that if shame did not contain a way of restoring severed or threatened ties it would be maladaptive because it would mean that to remain part of a group we could commit no error which is a demand that is virtually impossible to meet. Shame is however associated with behavioural tendencies that are similar to appeasement behaviours (Keltner et al. 1997, 360) These are apologetic and servile behaviours

shame in view, let's reconsider the case of Tony. There are two salient aspects of the case. First, Tony is a practicing Roman Catholic. Second, one of the tenants of Catholicism is that being gay is wrong. Proponents of DV are right here when they claim that Tony entertains a negative self-evaluation *of some sort* when he experiences shame. What happens is that Tony judges that he has failed according to the standards of his religious community. Although Tony does not accept those standards, in judging that he has failed according to them he entertains (but does not endorse) a negative self-evaluation. That evaluation when fully articulated would be something like "because of my attraction to other men, I am bad, according to Catholic standards" or "according to Catholic standards which I know to be wrong, I am bad because of my attraction to other men." What makes Tony's experience of shame recalcitrant is that a fitting evaluative judgment (e.g. being gay is wrong) is exactly the sort of evaluation that Tony *rejects*. Why on earth, we wonder, does Tony feel ashamed of being gay when he does not accept standards according to which being gay is wrong? The adaptive value of shame sheds light on this puzzlement. Tony's shame is, I believe, a result of the fact that his sexual preferences contradict one of the norms which delineate membership to the catholic community and his non-conformity to this norm acts as a barrier to his benefiting from the mutually co-operative relationships that exist within that group. In other words, Tony's shame is a reflection of the operation of the mechanisms which gave shame its adaptive value. At worst then, Tony's recalcitrant shame is not subject to rational scrutiny, because it is reflective of an engrained system which evolved for its adaptive benefits. At best, however, Tony's shame is rational because it is an instance of the successful operation of a system which helps us social animals to secure the benefits of group membership.

3. Vicarious Shame and Shame by Association

While the existence of recalcitrant shame puts pressure on DV, it does not force us to abandon it. The case of Tony, for example, shows that a subject can experience recalcitrant shame even if he does not (a) make a global and punitive judgment or (b) accept any criteria on the basis of which he would be negatively evaluated but it does not show that shame can be experienced in the absence of any negative self-evaluation. Tony does after all judge that he has failed according to the standards of his religious community and that is a negative self-evaluation of sorts. To show that DV should be abandoned, we need to

that "prevents or reduces aggression in others, results in social approach, and re-establishes severed or threatened social ties." (Maibom 2010 p. 578) Thus the behaviours associated with shame enable the shamed subject to be forgiven by and re-introduced into the group.

introduce a new range of cases: vicarious shame and shame by association. In both vicarious shame and shame by association a subject experiences shame for the wrong doings of others. A subject does not make a negative self-evaluation in either of these cases because the focus of the evaluations there is not *the subject* but rather an imagined or associated other. To begin our discussion of how these cases undermine DV, it is useful to first get clear of what vicarious shame and shame by association are, and why they are distinct from each other.

In shame by association, a subject experiences shame as a result of negatively evaluating the behaviour of a person or group that she is associated with. For example, if at a departmental party my partner gets horrifically drunk and I experience shame as a result, the shame that I experience is an instance of shame by association. That is to say, my shame is the result of my recognition of an association with a person whose conduct I evaluate negatively. Critically, it is my *recognition* of the association that generates this particular kind of shame experience. I would not feel shame by association if the person I witness behaving drunkenly at the departmental party is a stranger, not even if, as it turns out, that stranger is my long lost sister. In addition to my partner, I can be shamed because of my association to my family members, my political party, my school, work or union, my romantic partner or friend, my past, future or present government and my country. The range of these objects indicate that the thing to which a subject is associated when she is shamed by association may be connected to her through personal choice, e.g., her partner or political party, or it may not be, e.g., her family or her country. What is essential to shame by association is not that a subjects' relationship to the associated object is a product of personal choice but rather that she *recognises* her association to the object. In sum, shame by association is experienced as a result of the recognition of an association with someone or some group whose wrong doing you judge to be shameful.⁴⁷ Vicarious shame, on the other hand, is experienced as a result of empathetic perspective taking.⁴⁸ When I experience shame vicariously, I imagine myself in the role of another subject whose actions I take to be shameful. By adopting the others point of view, I experience shame through a recognition of their (imaginatively *my*) wrong doing. There is an embedded negative self-evaluation involved here (an evaluation of "my" wrong doing) but the subject of the evaluation is not *me* but rather the subject whose perspective I imaginatively adopt.⁴⁹ For this reason, in vicarious shame, I am required to keep track of two separate subjects, myself and the person I imagine myself to be. Importantly, the person whose perspective I imaginatively adopt need not in any way be associated to me.

⁴⁷ C.f. Walsh 1970

⁴⁸ Welten, Zeelenberg & Breugelmans 2012

⁴⁹ Martin 2006 p. 98

All that matters is that I am able to occupy their perspective and keep track of the fact that the occupied perspective provides a first personal take on a subject that is not *me*. Returning to the example of the departmental party, suppose it is a stranger whose drunken behaviour I witness, although I cannot experience shame by association, I can experience shame vicariously by imaginatively placing myself in their shoes. By taking on board the drunken person's perspective and from that perspective viewing their (imaginatively *my*) wrong doing I experience shame as if I was in that situation myself.

Both vicarious shame and shame by association are counter-examples to DV. In neither case am I required to make any kind of negative self-evaluation. In shame by association, I negatively evaluate an *associated object*. Here, it is not the self but something discernibly related to it that is the focus of evaluation. In some cases, it may be true that when I am shamed by association, I also judge that I am a failure or that I have failed according to some criteria because of my association with the object negatively evaluated. However, it seems to me at least bizarre to suppose that this kind of negative self-evaluation is *required* in all cases, particularly given the range of things one may be associated with when one is shame by association. Suppose I am shamed when I discover that Genghis Khan is a relation of mine. Why would I, or anyone else for that matter, be required to evaluate myself negatively because of the actions of a person whose relationship to myself I could not have chosen and whose actions I could have had no hope in influencing? In this case, it seems that the requirement for a negative self-evaluation is illegitimate precisely because an attribution of responsibility is inappropriate. If, however, one is not struck by the example of shame by association, the case of vicarious shame is even stronger. In cases of vicarious shame I recognise that I am responsible for a wrong doing and that recognition may be accompanied by the judgment that I am a failure or that I have failed according to some standards. However, in these cases I also hold in view that the subject of those negative self-evaluations, the person designated by the "I", is not *me* but rather *you* whose perspective I imaginatively adopt. That is to say, the subject of the negative evaluation embedded in vicarious shame and the subject of the shame experience come apart. As the subject of the negative evaluation involved in vicarious shame is the subject whose perspective I imaginatively adopt, it is not required for vicarious shame that I make any kind of negative *self*-evaluation.

I take it that neither shame by association nor vicarious shame requires a subject to make any kind of negative self-evaluation. If that is right, DV is false. Despite its faults however, DV at least gets right the centrality of the self to an experience of shame. We lose sight of what shame is if we sever it from a connection to the self and in particular, to a painful focus on the idea of the self. If shame by association and vicarious shame are to be counted

as instances of shame, we need an explanation of how it is that an evaluation of another persons' wrong doing is linked to the self and in particular to a painful feeling associated with the self. In other words, we need an explanation of why an evaluation of another persons' wrong doing can result in *shame*. Here again, the adaptive function of shame can prove useful. The adaptive function of shame is to track those of our behaviours that will be met with the threat of social exclusion. It is plausible to assume that the negative feeling to which shame is associated is derived from the negative consequences of that exclusion. The point, which shame by association and vicarious shame illustrates is, I think, that it is not just our own negative actions to which others may respond by excluding us from the group. The bad behaviours of others may also result in our exclusion from the group and as such the behaviours of others can be the cause of shame. Shame by association and vicarious shame are, I take it, a consequence of the recognition that the wrong doings of others may well result in our exclusion from the group. Why? In shame by association I identify with an associated object which is negatively evaluated. I identify with the associated object either because it is a group to which I belong (e.g., family, government, trade union) or because it is a person who bears some special relationship to me (e.g., my romantic partner). In identifying with the wrongdoer however, I recognise that their bad behaviour can reflect badly on me and by being tainted with their wrong doing; I run the risk of social exclusion. Hence I experience shame by association because the wrongdoings of the associated objects may result in my actual exclusion from the group. Things are different in the case of vicarious shame because in that case the risk of social exclusion is imagined. What happens here is that by putting myself in another person's shoes, I feel ashamed from the point of view of the person who has done the shameful deed. In other words, by taking on board the perspective of another in a shameful situation, I experience the imagined threat of social exclusion and it is this imagined threat which, I take it, causes vicarious shame.

4. Chapter Summary

To sum up: depending on how the notion of a negative self-evaluation is understood, three versions of DV are possible. I have tried to show that DV in all of its various forms fails to capture the multiplicity of our shame experiences. In recalcitrant shame a subject need not make a global punitive judgment and she need not negatively evaluate herself according to criteria that she accepts. That said, recalcitrant shame raised a puzzle. What was puzzling about recalcitrant shame was how we are to understand the fact that a subject who rejects a fitting evaluative judgment could experience shame without resorting to the claim that

this subject is irrational. In response to the puzzle, I argued that in some cases, recalcitrant shame is rational because in those cases, it displays a subjects' sensitivity to the values of the community to which she belongs and this sensitivity was essential to survival. Shame by association and vicarious shame provided an even stronger case against DV. In neither of these cases was it required that the subject makes any kind of negative self-evaluation. That said, it was unclear why a subject would feel shame as a result of the wrong doings of others. I gave a brief sketch of an account. In the case of shame by association, the suggestion was that I experience shame because the wrong doing of the associated other reflects badly on me and so may result in my exclusion from the group. In vicarious shame the suggestion was that because I imagine what it would be like to be in this shameful person's situation, I experienced an imagined threat of social exclusion. Here it is the imagined threat of exclusion which is responsible for my experience of shame. Of course, the examples of recalcitrant shame, shame by association and vicarious shame are not meant to show that negative self-evaluations are not a crucial part of some of our shame experiences. In many cases of shame, a subject will make a negative self-evaluation of some sort principally because the object of her shame will be a property of hers which she judges negatively. My point has only been to show that the requirement for a negative self-evaluation in all cases of shame cannot be right. Now, given the different forms a shame experience can take and the complex ways that these experiences can be related to the self, to social exclusion and to negative self-evaluations, we might wonder if anything at all general can be said about the causal mechanisms which generate a shame response. What is it about an event that causes an experience of shame? That is, after all the question that AT set out, unsuccessfully, to answer. In the next section I utilise Hume's model of the indirect passions to provide a hazy but overarching account of the causes of shame.

Chapter 3

1. Hume's Psychological Account of Shame

Shame is one of the indirect passions which Hume attempts to account for in his *Treatise on Human Nature*.⁵⁰ The passions are simple feelings.⁵¹ Being feelings of a certain sort, the passions lack representational content and so a constitutive or conceptual account of any one of them is not possible. Hume's view of the metaphysics of the passions shifts his attention away from giving a conceptual or constitutive analysis of the passions and towards giving an account of the circumstances in which the passions are aroused. Like an appraisal theorists then, Hume is interested in giving an account of emotional elicitation.⁵² Hume's account of the indirect passions comes in book II of the *Treatise*. Book II begins by recalling the distinction between original impressions and secondary impressions. Original impressions are internal sensations of physical pleasure or pain caused by perceptions of objects outside of ourselves. Hume claims that it is not possible to give an account of the causes of an original impression: it is just a brute fact that certain external objects generate certain bodily pleasures or pains.⁵³ While the cause of an original impression cannot be known, the cause of a secondary impression *can* be accounted for. Secondary impressions proceed either from reflection on original impressions or from reflection on the idea of an original impression. All of the passions are secondary impressions.

The passions are either direct or indirect.⁵⁴ The direct passions include hope, fear and grief. These "arise immediately from [reflection on] good or evil, from pain or pleasure."⁵⁵ The indirect passions include pride, shame, love and hatred. These arise in the same way as the direct passions but in conjunction with "other qualities".⁵⁶ Later, we learn that these additional qualities are in fact additional principles which are uniquely required to explain the elicitation of the indirect passions. The most important of those principles are the principles for the association of impressions and the principle of sympathy.⁵⁷ The principle of sympathy describes the circumstances in which one acquires an emotional state *x* on the basis of the belief that another subject is experiencing *x*. For now, we need only focus on the principle for the association of impressions. In book I of the *Treatise*

⁵⁰ It is generally accepted that what Hume says about Humility makes better sense if we suppose that for Hume, humility just is what we would call shame. C.f. Árdal 1966 p. 43

⁵¹ Hume T2.1.5.4

⁵² Hume T2.1.2

⁵³ Martin 2006 p. 75

⁵⁴ McIntyre 2000 p. 78

⁵⁵ Hume T2.1.4

⁵⁶ Hume T2.1.4

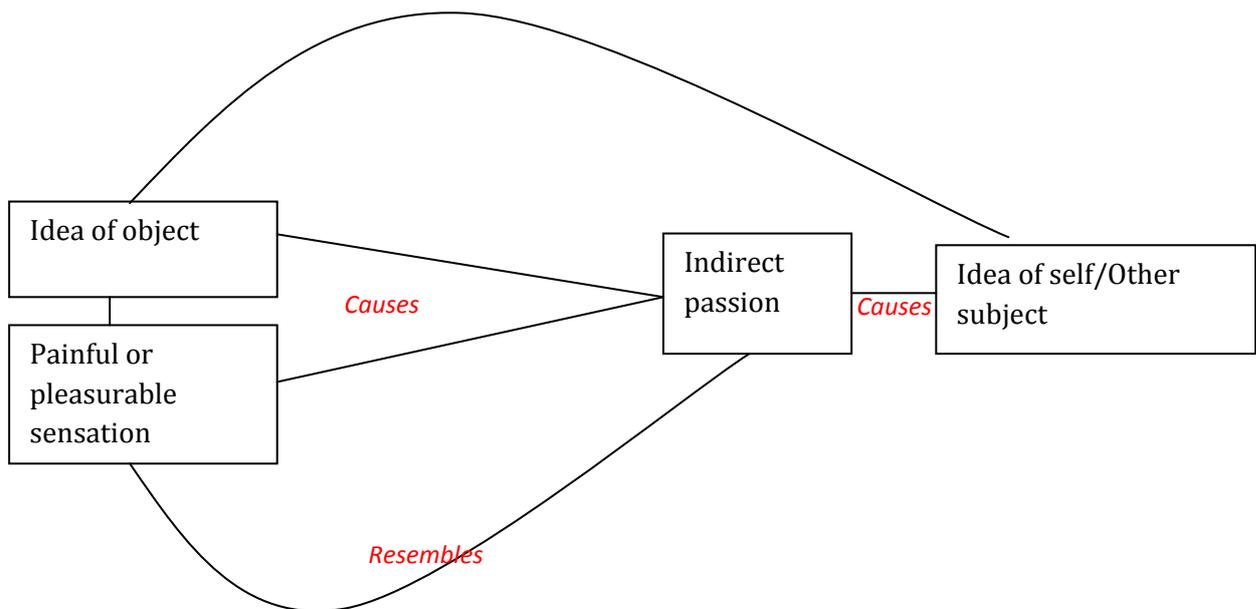
⁵⁷ Hume T2.1.11

Hume argues that ideas are naturally related by means of three associative principles: resemblance, contiguity and causation. In book II, the three principles governing the association of ideas are supplemented by a principle of association of impressions. Impressions, Hume tells us, are naturally associated by the relation of resemblance. For example the perception of a disagreeable impression will naturally lead us to the perception of another impression which resembles the first in that it is disagreeable. Thus we have three principles for associating ideas and one principle for associating impressions.

The principles of association are needed to explain the various interactions of the four elements that play a role in Hume's explanation of the causes of the indirect passions. These four elements are (1) the cause of the passion (an idea reflected on), (2) a positive or negative impression associated with the cause (a feeling of pleasure or pain), (3) the passion itself (a simple feeling) and (4) the object of the passion (what the passion is about). Features (1) and (2) can be grouped together as these are the joint causes of an indirect passion. So the cause of an indirect passion, on this view, is not merely an idea (i.e., a mental representation of an item of experience) but also a positive or negative quality to which that idea is associated (i.e., a *feeling* that is associated with the idea). Features (3) and (4) can be grouped together as the twin features of a passion because although Hume thinks that the passions are simple feelings, he thinks that an experience of an indirect passion causes the subject to think about the passion's object. The object of an indirect passion is either the idea of the self or the idea of another subject. It is, according to Hume, just a brute fact about human psychology that the experience of an indirect passion will cause a subject to reflect on the idea of herself or the idea of another subject.⁵⁸ To explain the elicitation of the indirect passions, all four elements are required. The basic idea is that the passion and its object are related to the idea and its property in a way which allows for an easy transition in the imagination from the cause to an experience of the passion.⁵⁹ More specifically Hume will suggest that reflection on an idea leads to an experience of an indirect passion when (a) that idea is related causally to the idea of the self or to the idea of another subject (i.e., the object of an indirect passion) and (b) the pleasurable or painful feeling associated with that idea is related via resemblance to the feeling which is the essence of the passion. We can represent this diagrammatically as follows:

⁵⁸ Hume T2.1.2.4

⁵⁹ Inoue 2003



Hume begins with preconceived ideas about whether an indirect passion is a pleasurable feeling or a painful feeling and whether the object of the passion is the idea of the self or the idea of another. It is just a basic fact about Hume's system that pride is a pleasurable feeling that has the self as its object and that shame is a painful feeling that has the self as its object. To explain how shame is elicited, Hume appeals to the structure above illustrated. Taking as basic the idea that shame is an unpleasant feeling with the self as its object, the diagram above tells us that any idea that has a quality which produces an unpleasant sensation can cause the unpleasant sensation that is distinctive of shame *so long as* the idea of the object is related by the principles of association to the object of shame (i.e., the idea of the self). The cause of shame therefore relates to the self and to an unpleasant sensation while shame relates to the self and *is* an unpleasant sensation. For Hume the principles of association exploit this parallel making it possible for a consideration of any idea that is related to the self and associated with a negative feeling to lead to an experience of shame. More exactly, the combined efforts of the principles of association for ideas and the principles of association for impressions link shame to the idea of the self, the negative feeling associated with the cause of shame to the negative feeling that is the essence of shame and the idea that is the cause of shame to the idea that is the object of shame. These various associations allow the mind to move easily and naturally from reflection on certain ideas to the experience of shame because that idea has features which parallel the features of shame.

In sum, Hume's thought is that reflection on an idea will cause shame when (a) the idea is associated with an unpleasant sensation which resembles the feeling of shame and (b) the idea is related to the self which is the object of shame. It is interesting to note how on this

model, Hume can accommodate the intentionality of shame (and of the indirect passions generally) despite his view that the passions are simple feelings. In essence the idea is that because the passions cause a subject to think about herself or another, the distinct sort of negative or positive feeling with which the passion is identified becomes associated with the idea of the self or the idea of another, in such a way that we can think of an indirect passion as a distinctive sort of pleasure or displeasure taken in the idea of the self or other. For example, in shame the idea of the self is associated with a painful feeling and as such, although shame is essentially a sort of negative feeling, it is a negative feeling directed at the idea of the self. Shame amounts to, on this view, a distinctive sort of displeasure taken in the idea of the self. Hume therefore accommodates the intentionality of shame through his requirement that the experience of shame *causes* a subject to reflect on herself, and in particular, to be pained by such self-reflection. It is the way in which the principles of association work to unite the four elements required to explain the cause of an indirect passion which allows Hume to explain the intentionality of the indirect passions. This is presumably why Martin (2006 p. 71) says:

“If we leave aside the details of Hume’s account because it requires us to talk of impressions and ideas and associationist principles, then we simply miss the distinctive ways in which Hume himself wishes to account for the directedness of the passions.”

In addition to accounting for the directedness of shame through the operation of the principles of association, Hume also seems to acknowledge that the idea, *x*, which causes shame does so in virtue of two beliefs that a subject has regarding it, namely the belief that *x* is associated with a negative sensation and the belief that *x* is related to me. As the cause of shame on Hume’s account is not merely an idea but a set of beliefs relating to that idea, Hume’s view appears to resonate well with cognitive theories of emotional elicitation. According to these theories thoughts (i.e., propositionally structured mental states) play an indispensable role in explaining the cause of our emotional experiences. We should, however, be cautious of the view that Hume is in fact providing a cognitive theory of emotional elicitation simply because beliefs play an indispensable role in Hume’s account of the indirect passions. For Hume beliefs are more akin to perceptions than to propositionally structured states so it would be controversial to pin onto Hume anything that we would recognise as a *cognitive* theory of the emotions.⁶⁰ Setting aside the question then of whether Hume *actually* provides a cognitive theory of emotional causation, some theorist have argued that with suitable modifications of Hume’s account (in particular by abandoning Hume’s atomism), we may arrive at a recognizably cognitive theory of

⁶⁰ Gorman 1993

emotions which preserves the spirit of Hume's view. For these theorists Hume should be praised because his view so closely approximates the correct account of the emotions, particularly of shame and pride. In the following sections, I wish to discuss the modifications of Hume's account which would have to be made in order to align his actual view of shame with a version of a cognitive theory of shame.

2. The Moral to Draw from Hume's Theory of Pride and Shame

In his (1989) paper, Árdal provides a reconstruction of Hume's theory of pride. Árdal begins his paper by agreeing with Davidson that "Hume's claim that pride is a simple impression is an embarrassment to him"⁶¹ and that "much of what he [Hume] says suggests the more plausible view that pride is essentially a form of self-evaluation."⁶² Árdal goes on to offer a reconstruction of Hume's view of pride according to which pride is "a favourable evaluation of oneself for a reason."⁶³ Given that shame is thought of by Hume as simply the opposite of pride⁶⁴ we can presume that Árdal would endorse a parallel reconstruction of Hume's view of shame, according to which shame is an unfavourable evaluation of oneself for a reason. We arrive at the reconstructed view of pride (and shame) that Árdal endorses by agreeing to a series of modifications to Hume's account. First, Árdal gets rid of Hume's atomism. The thought that an idea can be associated with a negative or positive impression is no longer taken to mean that the idea is a source of pleasure or displeasure for that subject but rather that a subject sees that object as having or lacking value.⁶⁵ So understood, Hume's claim that in shame the idea of the self is a source of displeasure is transformed into the claim that in shame a subject sees herself as lacking value. Similarly, Hume's claim that the cause of shame is viewed by the subject as a source of displeasure is transformed into the claim that the cause of shame is seen by the subject as lacking value. The causal story then becomes: if a cause is believed by a subject to lack value, and if it is believed to be related to herself, then this will cause her to feel shame and in feeling shame a subject is in a state of mind in which she thinks of herself as lacking value.

The second modification which Árdal suggests requires us to reconfigure Hume's understanding of the claim that shame has the self as its object as meaning that the experience of shame causes a subject to think about herself. Instead, we are to understand

⁶¹ Árdal 1989 p. 388

⁶² Árdal 1989 p. 388

⁶³ Árdal 1989 p. 389

⁶⁴ Hume T2.1.2.4

⁶⁵ Árdal 1989 p. 388

the claim that shame has the self as its object as meaning that shame *is* a way of negatively evaluating oneself. How do we get from the causal to the constitutive claim? We subtract from Hume the “embarrassing” idea that shame is a simple impression. Recall that for Hume, shame is a *simple* feeling. It is a simple feeling in the sense that it is not composed of any parts, it is not, for example a blend of two distinct feelings. Being a feeling that is not composed of parts Hume claims that a constitutive or conceptual account of shame is not possible. The most we can do, he thinks, is to determine the range of circumstances in which shame is caused and in this way to give an account of the causes of shame.⁶⁶ Now, by means of the first modification we were able to recast Hume’s claim that shame is a state in which thoughts about the self are a source of displeasure as the claim that in shame one negatively evaluates oneself. Now by disregarding Hume’s claim that shame is a simple impression we are able to give a constitutive account of shame after all. We can now say that shame *is* a state in which one negatively evaluates oneself or that shame *is* a way of negatively evaluating oneself. The final feature of Árdal’s modified account worth noting here is that the beliefs associated with the idea that is the cause of shame are transformed into *reasons* for a subject to be in that state.⁶⁷ Thus, we depart from Hume’s view of shame and arrive at the modified view according to which shame *is* a negative evaluation of the self *for* a reason. That reason is constituted by a subject’s belief that the cause of shame (the object of her idea) lacks value and her beliefs about the relationship between that object and herself.

Taking stock then, according to Árdal what we are supposed to learn from Hume is that shame is a form of self-evaluation and that the reason why a subject evaluates herself negatively in shame is because she believes that the cause of shame is negative and that it is related to her in some way. Clearly this view is not the one Hume advocates, but it has its roots in Hume and Árdal thinks that it depicts more accurately what shame actually is. What is right about Árdal’s modified account is that it requires that the connection between shame and the self be constitutive rather than causal. It is simply not possible to be in a state of shame that is not in some way connected to an idea about the self or a threat to the self or to ruminative thoughts about the self or to a perception of the self and so on. Where Árdal goes wrong, I believe, is with his introduction of self-evaluations into the Humean picture. The problem is that by aligning Hume’s view of shame with a version of DV, we instantly disallow Hume the ability to account for vicarious shame, shame by association and recalcitrant shame. I want to suggest that *as it stands*, Hume’s actual account of shame is better placed to explain the variants of shame that DV cannot. Let’s begin with shame by association.

⁶⁶ Hume T2.1.2.1

⁶⁷ Árdal 1989 p. 388-9

Suppose you experience shame in virtue of your association with a partner, whose drunken activities at the departmental party you judge to be shameful. How would Hume explain the elicitation of shame in this sort of case? Recall that for Hume the proximal cause of shame is not merely an idea but two beliefs relating to it. First, there is the belief that the idea is associated with a negative impression (i.e., that it is a source of displeasure) and second there is the belief that the idea is related to me in some way. Now there are two sorts of ways in which an object may be related to me. First, it may be related to me *directly* by being one of *my* actions or one of *my* physical or mental characteristics. Second, it may be related to me *indirectly* by being something to which I am associated (e.g., my boyfriend, my political party, my country and so on). The distinction between direct and indirect relations to the self is sketchy, but it will suffice for present purposes. We can see clearly that relations of an *indirect* kind are relevant when it comes to accounting for shame by association. When we factor this into Hume's account we get the following explanation: I reflect on an idea of my partner and given my beliefs about my relationship to him, thinking about him draws my attention to myself and given the unpleasant impression associated with that idea (i.e., his bad behaviour), this painful impression carries over to the idea of myself. I experience shame then because the unpleasant impression associated with my partner resembles the unpleasant sensation that is the essence of shame and the idea of my partner is related to the idea of myself which is the object of shame. The joint operation of the principles for the association of impressions and ideas means that reflection on my partner's bad behaviour at the party leads by an easy transition in the imagination to an experience of shame (i.e., a distinctively painful feeling connected to my idea of myself). With Hume's account then we can explain the elicitation of shame by association while preserving the critical connection between shame and the self.

If the moral to be drawn from Hume's theory is that shame is a form of self-evaluation, Hume would have a problem accounting for shame by association. The central difficulty would be to motivate the view that a negative self-evaluation is required to experience this sort of shame. What would ground such an evaluation? How is it that the belief that an object has a negative property and that it is indirectly related to me provides me with any reason to negatively evaluate myself? Árdall indirectly acknowledges the problem of grounds when he writes:

“The valuable characteristic that makes one proud is only sometimes a quality of the object of pride, the proud person himself. It is only when the quality belongs to the proud person that being proud seems to involve considering oneself praiseworthy. One may be proud of one's parents, ancestors, or country without

claiming to be entitled to praise... I hope to have made it clear that when people are proud of other things than themselves, they cannot claim to be praiseworthy as an aspect of their pride... Pride as an evaluation of the proud person is an evaluation of a kind that is quite different from the judgment that one is praiseworthy. One can legitimately claim to be proud of a related subject, although one had neither a hand in realizing the valuable cause nor a close relation to the cause, the two joint causes of pride." (p. 391-2)

From what I understand of this passage, it makes four key claims: (1) pride will involve an evaluation of oneself as praiseworthy only when the object of pride is a quality of the proud person (2) When I am proud by association (e.g. proud of my parents, my ancestors or my country) it would be illegitimate to evaluate *myself* as praiseworthy and so (3) the evaluation of the self required to experience pride (or at least pride by association) is of a different form to the evaluative judgment that I am praiseworthy. Finally (4) pride by association does not require the belief that the associated object is closely related to you or that you are in anyway responsible for realising the praiseworthy property in the associated object. As we are concerned here with shame, let's see what the corresponding claims, related to shame, would look like. (1*) shame involves an evaluation of the self as being unworthy of praise only when the object of shame is a quality of the subject who experiences that emotion (2*) When I am shamed by association it would be illegitimate to evaluate *myself* as being unworthy of praise and so (3*) the evaluation of the self required to experience shame (or at least shame by association) is of a different form to the evaluative judgment that I am not worthy of praise. Finally, (4*) shame by association does not require the belief that the associated object is closely related to you or that you are in anyway responsible for realising the property in the associated object that is not worthy of praise. At this stage, two questions become important: first, if the self-evaluations involved in shame by association are not of the sort that one is unworthy of praise, of what sort are they? Second, if it is not the belief that I am directly related to the object of shame or the belief that I am responsible for bringing about that objects negative characteristic, what grounds the self-evaluations involved in these cases?

The first question, *the kind question* is important because "not being worthy of praise" seems to amount to "not being worthy of positive evaluation". If that is the case there would be no other kind of negative self-evaluation that shame by association can involve because the judgement that one is not worthy of praise will be broad enough to cover any negative self-evaluation that one might care to suggest. This would force us to say that shame by association requires a negative evaluation which *entails* the judgment that one is not worthy of praise and that is precisely what we do not want. The second question, *the*

grounding question, is also important. If we cannot determine the grounds on the basis of which a self-evaluation is made then a subject who is shamed by association looks to be irrational. This is because a subject who is shamed by association will have no reason to accept the negative self-evaluation that she is required to make when she is in that experiential state. As Hume has no problem accounting for shame by association, Árdal's modified version of Hume's theory can be seen to be an improvement to the original only if he can adequately answer the grounding question and the kind question. It is therefore surprising that Árdal only addresses the grounding question in his paper. He writes:

“Is it not fundamentally irrational, however, to allow people to *claim* any kind of increase in value from a situation that does not involve any praiseworthy characteristic being possessed by the proud person? What constitutes a reason for pride differs from the reasons that justify praise. The proud person *in some way* acquires positive value by association with something valuable; *some of the value of the cause is allowed to rub off upon the object of pride*, the proud person.” (Árdal, 1989 p. 393 italics added)

The suggestion is that it is not irrational for a person to experience pride as a form of positive self-evaluation even when the cause of pride is not a characteristic that makes her worthy of praise. This is because what grounds the positive self-evaluation constitutive of pride need not be the same as what grounds an evaluation of oneself as praiseworthy. Indeed, in cases in which one is proud by association these two grounds part company. What grounds (and therefore *rationalises* the presence of) the positive self-evaluation constitutive of pride in cases in which one is proud by association is, according to Árdal, that the value associated with the object of pride is allowed *to rub off* onto the proud person. The suggestion, therefore, is that the appropriation of value from the associated object grounds (and therefore *rationalises* the presence of) the positive self-evaluation constitutive of pride. To assess this suggestion, we need to understand what it means for the value of an object associated with me to rub off onto me. Unfortunately, Árdal gives us no indication of how we are to understand the metaphor of “rubbing off” in this context and it seems to me that the most natural way of understanding this metaphor fails to give us an adequate answer to the grounding question

Let's consider a concrete example. Suppose that I am proud of my son's intelligence. What would it mean for some of the value associated with my son (i.e., his intelligence) to rub off onto me? One thing it might mean is that I would myself come to possess, or believe that I possess, some intelligence in virtue of my relationship to him. This interpretation is suggested by the ordinary meaning of the phrase “rub off on” which is to pass on a quality

or characteristic. But this interpretation cannot be right. I am not and do not believe myself to be intelligent, not even a little bit intelligent, in virtue of my association with an intelligent person. The same worry arises when the property of the related object is not itself a value but rather something judged to be *valuable*. For example, suppose that I am proud of my son because he produces beautiful paintings. My son's paintings are not values but they are valuable because they are beautiful. Árdal would claim that in this case, I am required to positively evaluate myself and that what grounds my positive self-evaluation is that some of the value of the associated object is allowed to rub off onto me. As the value which I acquire in this case is (some amount of) the property in virtue of which my son's painting is valuable we arrive at exactly the same problem. We would be forced to accept the implausible claim that I am or believe myself to be beautiful because of my association with the person that produced these beautiful paintings and that it is this fact or belief that grounds (and therefore rationalises) the positive self-evaluation that I am required to make. The problem here is simply that it makes little sense to talk about acquiring a value x (or some share of that value) simply in virtue of being associated with a person that exemplifies x or who produces an artefact that exemplifies x . Values aren't contagious in the relevant sense.

How then are we to understand the idea of values rubbing off onto the associated object? It is unclear, however some account is needed here because without an answer to the grounding question, the view that shame requires a negative self-evaluation is no improvement on Hume's view.⁶⁸ This is because Hume's account as it stands can easily explain shame by association *and* it can readily explain those cases in which shame does involve a negative self-evaluation. Let me explain. In Hume's explanation of the causes of shame, we can distinguish two ways in which an object can be related to the subject: directly or indirectly. When the object of shame (the idea which causes it) is directly related to the subject it will be a property of the subject herself. Shame caused by that object will be related in a straightforward way to a negative self-evaluation because when the cause of shame is a personal property that is a source of pain, it is easy to see why the subject of that state would negatively evaluate herself. In other words, it is easy to see why viewing a personal property negatively would provide grounds for a negative self-evaluation and as such we can expect that shame will involve a negative self-evaluation when the object of shame is directly related to one. When the cause of shame is not a personal property that is a source of pain but an associated object which is a source of

⁶⁸ I suggested (p. 52 above) that because of the range of things to which one can be associated when one is ashamed by association we cannot think that being associated with an object evaluated negatively *in all cases* provides one with a reason to negatively evaluate oneself.

pain, it becomes puzzling to see why a negative self-evaluation is required simply because it is unclear what would ground such an evaluation.

Thus, without any modification of his account whatsoever, Hume can adequately explain shame by association and, what I will call, ordinary shame: cases in which the cause of shame is a personal property viewed negatively. In these latter cases, shame is likely to involve a negative self-evaluation because the negative viewing of a personal attribute provides excellent grounds for a negative self-evaluation and presumably, a rational subject will recognise those grounds and be moved accordingly to make a negative self-evaluation. In addition to explaining the elicitation of shame by association and ordinary shame, Hume's account can also be used to explain the elicitation of recalcitrant shame. In recalcitrant shame, a subject's shame is caused by an object which she does not negatively evaluate. For example, Tony experienced recalcitrant shame as a result of his sexuality even though he did not think that there was anything wrong with being gay. Hume can explain Tony style cases in exactly the same way as he does ordinary shame and shame by association. The idea is that the cause of shame (the idea of being gay) is related to the subject and is associated with an unpleasant feeling. This unpleasant feeling resembles shame, and because the cause of shame is itself related to the object of shame, the principles of association of impressions and ideas renders the transference from a consideration of the cause (the idea of being gay) to the experience of shame easy and natural. What might now appear puzzling is why it is that, for Tony, the idea of being gay is associated with an unpleasant feeling. However, there is no reason to suppose that the association between an idea and an impression cannot in some cases be idiosyncratic. That is to say, there is no reason to suppose that the connection between a particular idea and a particular impression cannot be forged simply because of the subjects own personal history and experience. If, however, a more general explanation is required, we may appeal to the evolutionary story told above. The idea would be that being gay is a source of pain for Tony because he is aware that others in his social group may view him negatively in virtue of his sexuality.

3. A Special Look at Hume on Vicarious Shame.

Hume can account for recalcitrant shame, ordinary shame and shame by association however the task of accounting for vicarious shame is for Hume, that much harder. That said, I believe that Hume's system can go some way to explaining the elicitation of vicarious shame. To see how Hume is able to do this, we need to revisit the second of the two additional principles required to explain the elicitation of the indirect passions,

namely, the principle of sympathy. Sympathy is not an emotion on Hume's account, but rather a principle that is invoked to explain the transformation of an idea of someone's passion into the experience of that passion.⁶⁹ What makes it possible for an idea of your passion to transform into an experience of that passion is Hume's theory of ideas and impressions. According to Hume's theory of ideas and impressions, the only difference between an idea and an impression is the level of force and vivacity with which they strike the mind: an impression appears more vividly than an idea does.⁷⁰ As the passions are all secondary impressions, all that is required for my idea of your passion to turn into the passion itself is an increase in the level of force and vivacity with which that idea strikes my mind. The principle of sympathy then is invoked to explain how it is that my idea of your passion can increase in force and vivacity and therefore transform into my experience of that passion. The first thing to note about sympathy is that it can only take effect when the sympathiser and the person sympathised with are related by three associative principles. These are the principles of similarity, the principles of contiguity in time and space and the principles of cause and effect.⁷¹ Why exactly must I be related to you by these three associative principles in order for my idea of your passion to transform into the passion itself?

The relation of cause and effect is primarily required to make an inference from the cause or effect of the passion to the passion itself. That is to say, it is required so that the sympathizer can infer from the behaviour or circumstance of the person sympathised with that she is in a certain emotional state. Through the use of the principles of cause and effect then, the sympathizer arrives at the idea of the other persons' passion (i.e., she arrives at a belief about the emotion that the other person is feeling). The relation of similarity is required because, according to Hume, it is only by being similar to the person that it is possible for me to come to be in the emotional state that they are in.⁷² Hume thinks that there is a natural resemblance between all humans which makes it possible for any one person to sympathise with any other person,⁷³ however he notes that it is easier to sympathise with those people that are similar to us in more particular respects, for example, a fellow countryman.⁷⁴ The relation of contiguity in time and space is required for sympathy to take effect principally because it facilitates the ease with which the mind moves from the idea of the others passion to the experience of the passion.⁷⁵ Generally, we sympathize with our neighbours more easily than with people who live in other parts of

⁶⁹ Hume T2.1.11

⁷⁰ Hume T1.1.1

⁷¹ Hume T2.1.11

⁷² Hume T2.1.11

⁷³ Hume T2.1.11, T2.2

⁷⁴ Hume T2.1.1

⁷⁵ Hume T2.3.7

the world and we sympathise with people from our own time more easily than we do with people from distant times. According to Hume, this is because the relation of contiguity in time and space makes the movement of the mind required from the idea to the passion more easy and natural. In sum, the three principles of cause and effect, similarity and contiguity are each necessary for sympathy to take effect: the principle of cause and effect is required to arrive at a belief about the passion that the other person is experiencing based on an inference from their behaviour or their circumstances, the principle of similarity is required in order for my idea of someone's passion to transform into the passion itself and the principle of contiguity is required to make this transference from the idea to passion more easy and natural. But *how* exactly does the transformation from idea to passion work? Hume explains:

'Tis evident, that the idea, or rather impression of ourselves is always intimately present with us, and that our consciousness gives us so lively a conception of our own person, that 'tis not possible to imagine, that any thing can in this particular go beyond it. Whatever object, therefore, is related to ourselves must be conceived with a like vivacity of conception according to the foregoing principles... (T 2.1.11, 317).

Hume's idea here is that we have a lively impression of ourselves which is "always intimately present with us." The impression of the self which Hume takes to be forever in our purview is constituted by a succession of perceptions.⁷⁶ When I am related to you via the principles of resemblance, causation and contiguity, my impression of myself enlivens the idea of your passion and this is what transforms my idea of your passion into the passion itself. In other words the liveliness of our conception of ourselves is transferred to the idea of the other's passion, making the idea of their passion so lively that it becomes the passion itself.⁷⁷ But according to Hume, this transference can occur only when we are related to the person believed to experience the passion by the relations of resemblance, contiguity, and cause and effect. This is because the more strongly I am related to you via these relations, the greater the chance that my conception of myself will enliven my idea of your passion and so transform this idea into a passion.⁷⁸ From all of this we can see that Hume has the resources to account for something like emotional contagion, which is a very strong element of what is involved in vicarious shame. In vicarious shame a subject derives from another's situation the belief that the other is or is likely to be experiencing emotion *x* and then she *somehow* transforms her belief about the others emotional state

⁷⁶ Hume T2.1.2, T1.4.6

⁷⁷ Herdt 1997 p. 39-40

⁷⁸ Hume T2.1.11

into an experience of that emotion. Hume gives us an explanation of *one way* in which a belief about another's actual or likely emotional state can be transformed into the actual sentiment and so he goes some way to explaining the elicitation of vicarious shame. However, there is a problem here. Hume cannot be said to fully account for vicarious shame simply through the operation of the principle of sympathy because his account of the self as a succession of perceptions makes it difficult to see how one could occupy the perspective of another subject. That is to say, if the self is just a bundle of successive perceptions, it becomes difficult to see how it is that I can put myself in the mind of another subject. What Hume can explain clearly with the principle of sympathy is the transference of an emotion from person *x* to person *y*. What Hume cannot seem to explain, given his conception of the self, is how the transference of an emotion from person *x* to person *y* can come about through one's imaginatively occupying another person's perspective.⁷⁹ It seems then that Hume's theory can only go some way to explaining the elicitation of vicarious shame but in this respect, it is superior to the dominant account of shame which cannot even recognise the possibility of vicarious shame.

4. Chapter Summary

In book II of the *Treatise* Hume offers an account of emotional elicitation for a particular range of emotions including shame and pride. It has been argued by some theorists that Hume has within his account the makings of a cognitive theory of pride and shame. The idea seems to be that, if we modify Hume's theory in such a way as to align it with a variant of DV, we will gain an improvement of Hume's account which nevertheless captures some of the flavour of Hume's actual account. I think we should resist the suggestion that Hume's view is better when it is modified in such a way as to align it with a version of DV. This is because Hume's actual theory provides a model which can be used to explain the elicitation of ordinary shame, shame by association, recalcitrant shame, and perhaps with some modifications of Hume's account of the self, vicarious shame. That is to say DV does not represent an improvement on Hume's account of shame but rather Hume's account of shame represents an improvement on DV.

Not that Hume's theory is perfect. Hume does not seem to distinguish between being in an emotional state and feeling emotion⁸⁰, nor does he seem to have the resources to explain what it is to imaginatively occupy the perspective of another.⁸¹ The purpose of this chapter however has not been to blindly accept a Humean view of shame but rather to

⁷⁹ Martin 2006 p. 97-98

⁸⁰ Árdal 1989 p. 25

⁸¹ Martin 2006 p. 97-98

demonstrate its remarkable flexibility. If, like Ardal, we interpret the moral of Hume's theory as one according to which shame involves a negative self-evaluation we rob Hume of the ability to account for a range of cases which proponents of DV cannot explain. Unlike Ardal then, I take it that the moral we are to draw from Hume's account is not that self-evaluations are constitutive of shame but rather that there is a relationship between the cause of shame and the subject of shame such that thinking about the cause puts that subject into a certain painful experiential state that is intimately related to a subjects view of herself.

Conclusion

The purpose of this essay has been to argue against DV. According to DV, a negative self-evaluation is required to experience shame. I began in chapter one by elucidating a causal variant of DV, which claimed that in order for a stimulus x to elicit a shame response that stimulus must prompt the subject to make a negative self-evaluation. What is interesting about this particular variant of DV is that some psychologists have attempted to provide empirical support for it. The main purpose of the first chapter was to assess how strong a case could be made for DV on the basis of empirical evidence. It was argued that of the three methods used by psychologists to establish conclusions about the causal relationship between appraisals and emotions only the elicitation method could establish both that there is a causal relationship between appraisals and emotions and that the direction of causality moves from appraisal to emotions. The problem however is that while evidence derived from the elicitation method may show that negative self-evaluations are sometimes required for an experience of shame they could not show that shame is not also, and perhaps frequently, experienced in cases in which no negative self-evaluation is made. In other words, although we may well use the elicitation method to provide empirical evidence for the claim that shame is sometimes caused by self-evaluations, this evidence *could not* be used to support the claim that self-evaluations are *required* to experience shame because there are cases of shame which do not involve negative self-evaluations; cases which it is the purpose of chapter two to discuss.

In chapter two, I raised three sorts of cases which versions of DV have a difficulty accounting for. These were recalcitrant shame, shame by association and vicarious shame. In recalcitrant shame a subject experiences shame directed at an object or cause x even when she is aware that her evaluation of x is at odds with her experience of shame. In shame by association a subject experiences shame in virtue of her association with another person or group whose actions she negatively evaluates. In vicarious shame, a subject experiences shame as a result of imaginatively adopting the perspective of another subject whose situation she judges to be worthy of shame. In each of these cases similar puzzles were raised. In the case of recalcitrant shame the puzzle was to understand why it is that a subject would experience shame even when she did not judge that the cause or object of her shame was shameful. In the case of shame by association the puzzle was to understand why it is that a subject would experience shame when it was not her but an associated subject whose conduct was evaluated as shameful. Finally, in the case of vicarious shame the puzzle was to understand why it is that a subject could experience shame when it is not her but a subject whose perspective she imaginatively adopts that

has behaved shamefully. In each of these cases, I take it that what is required is a rationalizing explanation of the shame experienced. In other words, we need to explain why given the particular circumstances in which vicarious shame, shame by association and recalcitrant shame are elicited a subject would feel shame at all. What emerged in chapter two was that the adaptive function of shame could be useful in providing a rationalising explanation of each of these types of shame experience.

The adaptive function of shame is to detect those behaviours that will be met by the threat of social exclusion. This threat is itself a source of fear and anxiety because being excluded from the group means being excluded from co-operative activities which promote and facilitate the individuals survival. The adaptive function of shame shows that it is a sufficient requirement to experience shame that one recognizes that some behaviour will lead to the threat of social exclusion and one feels fear or anxiety as a result of this threat. Note that it is the groups values which determine the things that are worthy of the punishment of exclusion. The subjects' own attitudes to these things are irrelevant, so it may be that the things which the group judges to be wrong and so as warranting exclusion are not the things that a subject herself judges to be wrong or as warranting exclusion. With the adaptive function of shame in view, I think it is possible to explain why given the particular circumstances in which vicarious shame, shame by association and recalcitrant shame are elicited a subject feels shame. In recalcitrant shame a subject feels shame because she detects that others in a group to which she belongs will negatively evaluate the object or cause of her shame and as a result will exclude her from the group because of her relationship to that cause or object. In shame by association a subject feels shame because she detects that in virtue of her relationship with the associated object (whose conduct is negatively evaluated) she would be threatened with exclusion from a group to which she is a member. In vicarious shame the subject feels shame because she imaginatively puts herself in the position of another subject and detects from this perspective that her behaviour would be met with the threat of social exclusion. Here it is the *imagined* threat of social exclusion as the result of an imagined wrongdoing which explains why the subject experiences shame. Thus, by appealing to the adaptive function of shame we can give a rationalizing explanation of the experience of shame in all of the three cases discussed.

I concluded from chapter two that DV is inadequate and that we need to look elsewhere to provide an overarching account of the causes of shame. In the final chapter I looked at Hume's account of shame. Hume, like the appraisal theorist, offers an account of the elicitation of shame. What is peculiar about Hume's account, however, is that no mention

is made of a negative self-evaluation. Hume's account of shame is set against his view that the principles of the mind, which include the principles of the association of impressions, the principles of the association of ideas and the principles of sympathy, make the transition between certain thoughts very natural and easy. In shame these principles operate in such a way as to make a certain range of ideas (the causes of shame) lead naturally to an experience of shame. The ideas which cause shame are both related to the self and are associated with a painful impression. In Hume's system it is a basic fact that shame is a painful feeling which is associated with the self. What happens is that because the idea which is the cause of shame is related to the self which is the object of shame and because the feeling associated with the idea is a painful impression, which resembles the experience of shame, the mind moves easily from a consideration of these ideas to the experience of shame. Using this account, Hume can explain shame by association, recalcitrant shame and ordinary shame but it is only by appealing to the operation of the principle of sympathy that Hume can go some way to explaining vicarious shame. It is because of the versatility of Hume's account that I chose to end the thesis with a discussion of it. Hume's theory provides an interesting point of departure for philosophical and psychological theorists looking for an overarching account of the elicitors of shame.

Bibliography

1. Árdal, P. S., (1966) *Passion and Value in Hume's Treatise*. Edinburgh University Press.
2. Árdal, P. S., (1989) 'Hume and Davidson on Pride'. *Hume Studies* 15 (2): 387-394.
3. Calhoun, C., (2004) 'An Apology for Moral Shame'. *Journal of Political Philosophy* 12 (2): 127-146
4. Baker, R. L., (2011) 'How to Have Self-Directed Attitudes' in Konzelmann Ziv, A., Lehrer, K. & Schmid, H. B., (Eds.) *Self-Evaluation – Affective and Social Grounds of Intentionality*. Philosophical Studies Series 116, Dordrecht: Springer
5. Bartky, S., (1990) *Femininity and Domination*. New York: Routledge.
6. Brady, S. M., (2009) 'The Irrationality of Recalcitrant Emotions'. *Philosophical Studies* 145 (3): 413 - 430.
7. D'Arms, J. & Jacobson, D., (2003) 'The Significance of Recalcitrant Emotion' in Hatzimoysis, A. (Ed.), *Philosophy and the Emotions*. Cambridge University Press.
8. Darwin, C., (1872) *The Expression of Emotion in Man and Animal*. London: John Murray.
9. Davidson, D., (1978) 'Hume's Cognitive Theory of Pride'. *Journal of Philosophy* 75 (19): 744-757
10. Davidson, D., (1980) *Actions and Events*. Oxford University Press.
11. Deonna, J. & Teroni, F., (2009) 'The Self of Shame' in Salmela, M. & Mayer, V., (Eds.), *Emotions, Ethics and Authenticity*. John Benjamins.
12. Deonna, J., Rafeale, R. & Teroni, F (2011a) *In Defense of Shame*. New York: Oxford University Press.
13. Deonna, J. & Teroni, F., (2011b) 'Is Shame a Social Emotion?' in Konzelmann Ziv, A., Lehrer, K. & Schmid, H. B., (Eds.) *Self-Evaluation – Affective and Social Grounds of Intentionality*. Philosophical Studies Series 116, Dordrecht: Springer.
14. Deonna, J., & Teroni, F., (2008) 'Distinguishing Shame from Guilt'. *Consciousness and Cognition* 17 (4): 1063-1400.
15. Dooren, M. V., De Vries J. J. G & Janssen, J. H., (2012) 'Emotional Sweating Across the Body: Comparing 16 Different Skin Conductance Measurement Locations'. *Physiology & Behavior* 106 (2): 298-304
16. Ekman, P., & Friesen, W. V., (1975) *Unmasking the Face: A Guide to Recognizing from Facial Clues*. New York: Prentice Hall.

17. Ekman, P., Davidson, R. J. & Friesen, W. V., (1990) 'The Duchenne Smile: Emotional Expression and Brain Physiology II'. *Journal of Personality and Social Psychology* 58, 342-353
18. Ekman, P., (1992a) 'Are There Basic Emotions?' *Psychological Review* 99 (3): 550-553
19. Ekman, P., (1992b). 'An Argument for Basic Emotions'. *Cognition and Emotion* 6, 169-200.
20. Ekman, P., (1999) 'Basic Emotions' in Dalglish, T. & Power, M., (Eds.) *Handbook of Cognition and Emotion*. John Wiley & Sons Ltd
21. Ekman, P., (2003) *Emotions Revealed*. New York: Times Books.
22. Fessler, D. M. T., (2004) 'Shame in Two Cultures: Implications for Evolutionary Approaches.' *Journal of Cognition and Culture* 4(2):207-262.
23. Frankfurt, H., (1988) *The Importance of What We Care About: Philosophical Essays*. Cambridge University Press.
24. Frijda, N. H., (1986) *The Emotions*. Cambridge University Press.
25. Frijda, N. H., (1988) 'The Laws of Emotion'. *American Psychologist* 43, 349-358.
26. Frijda, N. H., (2007) *The Laws of Emotions*. Lawrence Erlbaum Associates.
27. Frijda, N. H., & Zeelenberg, M., (2001) 'Appraisal: What is the Dependent?' in Scherer, K. R, Schorr, A. & Johnstone, T. (Eds.) *Appraisal Process in Emotion: Theory, Methods, Research*. Chicago: University of Chicago Press
28. Greenspan, P., (1988). *Emotions and Reasons*. London: Routledge
29. Griffith, P. E., (2003) 'Basic Emotions, Complex Emotions, Machiavellian Emotions' in Hatzimoysis, A. (Ed.) *Philosophy and the Emotions*. Cambridge University Press.
30. Hatzimoysis, A., (2003) *Philosophy and the Emotions*. Cambridge University Press.
31. Helm, B., (2001) *Emotional Reason*. Cambridge University Press
32. Herdt, J. A., (1997) *Religion and Faction in Hume's Moral Philosophy*. Cambridge Studies in Religion and Critical Thought 3. Cambridge University Press
33. Hume, D., *Treatise of Human Nature*. References to are by Book. Part. Section. Page references to the New Oxford Edition: Hume, D., (2000), *A Treatise of Human Nature*. Norton D.F. & M.J. Norton.
34. Inoue, H., (2003) 'The Origin of the Indirect Passions in the Treatise: An Analogy Between Books 1 and 2.' *Hume Studies* 29 (2): 205-22
35. Izard, C. E., (1991) *The Psychology of Emotions*. New York: Plenum.

36. Izard, C. E., (1993) 'Four Systems for Emotion Activation: Cognitive and Noncognitive Processes'. *Psychological Review* 100 (1): 68–90.
37. Izard, C. E., Ackerman, B. P. & Schultz, D., (1999). 'Independent Emotions and Consciousness: Self-consciousness and Dependent Emotions' in Singer, J. & Salovey, P., (Eds.) *At Play in the Fields of Consciousness: Essays in Honor of Jerome L. Singer*. Mahwah, NJ: Erlbaum.
38. Keltner, D. & Buswell, B. N., (1997). 'Embarrassment: Its Distinct Form and Appeasement Functions'. *Psychological Bulletin* 122, 250-270.
39. Kenny, A., (1963) *Action, Emotion, Will*. London: Routledge & Keegan Paul
40. Konzelmann Ziv, A., (2011) 'Self-Evaluations: Philosophical Perspectives' in Konzelmann Ziv, A., Lehrer, K. & Schmid, H. B., (Eds.) *Self-Evaluation – Affective and Social Grounds of Intentionality*. Philosophical Studies Series 116, Dordrecht: Springer.
41. Lazarus, R. & Alfert, E., (1964) 'Short-circuiting of Threat by Experimentally Altering Cognitive Appraisals'. *Journal of Abnormal and Social Psychology* 69, 195-205
42. Lazarus, R., (1991) *Emotion and Adaption*. Oxford University Press.
43. Lazarus, R., (2001) 'Relational Meanings and Discrete Emotions' in Scherer, K. R., Schorr, A. & Johnstone, T., (Eds.) *Appraisal Process in Emotion: Theory, Methods, Research*. Oxford University Press.
44. Leary, M. R., (2007) 'How the Self Became Involved in Affective Experience: Three Sources of Self-reflective Emotions' in Tracy, R. & Tangney, J., (Eds.) *The Self-Conscious Emotions: Theory and Research*. New York: Guilford.
45. Lewis, M., (2000) 'Self-conscious Emotions: Embarrassment, Pride, Shame, and Guilt' in Lewis, M. & Haviland-Jones, J. M., (Eds.) *Handbook of Emotions*. 2nd Edition New York: Guilford.
46. Lewis, M., (2003) 'The Role of the Self in Shame' *Social Research* 70 (4): 1181-1204
47. Lewis, M., (1999) 'The Role of the Self in Cognition and Emotion' in Dalgleish, T & Power, M., (eds). *Handbook of Cognition and Emotion*. John Wiley & Sons Ltd Publishers.
48. Lewis, H. B., (1971) *Shame and Guilt in Neurosis*. New York: International Universities Press.
49. Lewis, D., (1972) 'Psychophysical and Theoretical Identifications'. *Australasian Journal of Philosophy* 50, 249–58; reprinted in Rosenthal, D., (1994) All page references are to the reprint.

50. Martin, M. G. F., (2006) 'In Praise of Self: Hume's Love of Fame'. *European Journal of Analytic Philosophy* 2 (1): 69-100.
51. McEachrane, M., (2009) 'Emotion, Meaning and Appraisal theory' *Theory & Psychology* 19 (1): 22-53
52. McIntyre, J. L., (2000) 'Hume's Passions: Direct and Indirect', *Hume Studies* 26 (1): 77-86.
53. Niedenthal, P. M., Tangney, J. P. & Gavanski, I., (1994) "'If only I weren't" versus "If only I hadn't": Distinguishing Shame and Guilt in Counterfactual Thinking'. *Journal of Personality and Social Psychology*. 67(4): 585-595
54. Öhman, A. & Soares, J. J. F., (1994). 'Unconscious Anxiety: Phobic Responses to Masked Stimuli'. *Journal of Abnormal Psychology* 103, 231-240.
55. Öhman, A., (2000) 'Automaticity and the Amygdale: Nonconscious Responses to Emotional Faces.' *Current Directions in Psychological Science* 11, 62-66.
56. Parkinson, B., (1995) *Ideas and Realities of Emotion*. London: Routledge.
57. Roseman, I. J., (1991) 'Appraisal Determinants of Discrete Emotions'. *Cognition and Emotion* 5, 161-200
58. Roseman, I. J. & Smith, C. A., (2001) 'Appraisal theory: Overview, Assumptions, Varieties, Controversies' in Scherer, K. R., Schorr, A. & Johnstone, T., (Eds.) *Appraisal Process in Emotion: Theory, Methods, Research*. Oxford University Press.
59. Roseman, I. J. & Evdokas, A., (2004) 'Appraisals Cause Experienced Emotions: Experimental Evidence'. *Cognition and Emotion* 18, 1-28.
60. Rosenthal, D., (1994) *The Nature of Mind*. Oxford University Press.
61. Sartre, J. P., (1943/1957) *L'Être et le néant*. Éditions Gallimard. Paris.
62. Scherer, R. K., (1999) 'Appraisal theory' in Dalgleish, T. & Power, M. J., (Eds.), *Handbook of Emotion and Cognition*. Chichester: New York.
63. Scherer R. K., (2001) 'The Nature and Study of Appraisals: A Review of the Issues' in Scherer, R. K., Schorr, A., & Johnstone, T., (Eds.) *Appraisal Process in Emotion: Theory, Methods, Research*. Oxford University Press.
64. Schorr, A. (2001a). 'Subjective Measurement in Appraisal Research: Present State and Future Perspectives' in Scherer, R. K., Schorr, A., & Johnstone, T., (Eds.) *Appraisal Process in Emotion: Theory, Methods, Research*. Oxford University Press.
65. Schorr, A., (2001b) 'Appraisal: The Evolution of an Idea' in Scherer, R. K., Schorr, A. & Johnstone, T., (Eds.). *Appraisal Process in Emotion: Theory, Methods, Research*. Oxford University Press.
66. Smith, C. A. & Ellsworth, P. C., (1985) 'Patterns of Cognitive Appraisal in Emotion'. *Journal of Personality and Social Psychology* 52, 475-488.

67. Smith, C. A., and Lazarus, R., (1993) 'Appraisal Components, Core Relational Themes and the Emotions.' *Cognition and Emotion* 7 (3): 233-269
68. Solomon R. C., (2003) *Not Passion's Slave: Emotions and Choice*. Oxford University Press
69. Solomon, R. C., (1976) *The passions: Emotions and The Meaning of Life*, Indianapolis, IN: Hackett Publishing Company.
70. Speisman, J. C., Lazarus, R. S., Mordkoff, A. & Davison, L., (1964) 'Experimental Reduction of Stress Based on Ego Defence Theory'. *Journal of Abnormal and Social Psychology* 68, 367-80
71. Sroufe, A., (1995). *Emotional Development: The Organization of Emotional life in the Early Years*. New York: Cambridge University Press.
72. Taylor, G., (1985) *Pride, Shame, and Guilt Emotions of Self-assessment*. Oxford Oxfordshire: Clarendon Press.
73. Tracy, L. J, Robins, R & Tangney, P.J., (2007) *The Self-Conscious Emotions: Theory and Research*. The Guilford Press
74. Tracy, J. L & Robins, R., (2004) 'Putting the Self into Self-Conscious Emotions: A Theoretical Model'. *Psychological Inquiry*, 15 (2) 103-125
75. Velleman, D. J., (2006) *Self to Self: Selected Essays*. Cambridge University Press.
76. Walsh, W. H., (1970) 'Pride, Shame and Responsibility'. *The Philosophical Quarterly* 20 (78): 1-13
77. Weiner, B., (1985) 'Spontaneous' Causal Thinking'. *Psychological Bulletin* 97, 74-84
78. Welten, S. C. M, Zeelenberg, M. & Breugelmans, S. M., (2012) 'Vicarious Shame' *Cognition & Emotion*. DOI:10.1080/02699931.2011.625400
79. Wittgenstein, L., (2001) *Philosophical Investigations*. (G. E. M. Anscombe, Trans., 3rd edition). Malden, MA: Blackwell.
80. Zajonc, R. B., (1980) 'Feeling & Thinking: Preferences Need no Inference', *American Psychologist* 35, 151-75
81. Zajonc, R. B., (1984) 'On the Primacy of Affect'. *American Psychologist* 39, 117-123