Erkenn (2009) 71:123–139
DOI 10.1007/s10670-009-9171-0

ORIGINAL ARTICLE

Affective Ignorance

Christoph Jäger

Received: 10 March 2009 / Accepted: 10 March 2009 / Published online: 12 May 2009
© Springer Science+Business Media B.V. 2009

Abstract According to one of the most influential views in the philosophy of self-knowledge each person enjoys some special cognitive access to his or her own current mental states and episodes. This view faces two fundamental tasks. First, it must elucidate the general conceptual structure of apparent asymmetries between beliefs about one's own mind and beliefs about other minds. Second, it must demarcate the mental territory for which first-person-special-access claims can plausibly be maintained. Traditional candidates include sensations, experiences (of various kinds), thoughts, beliefs, desires, and also affective states such as emotions. I reconstruct five prominent privileged access claims that have traditionally been maintained for emotions and discuss logical relations among them. I then argue that none of these claims stands up to scrutiny. The truth is that we often suffer from affective ignorance, and that third-person ascriptions of emotional states should often be credited with more rather than less authority than corresponding self-ascriptions. I conclude by considering, and rejecting, five potential objections to my argument.

1 Introduction

Here is a sample of statements in which prominent philosophers and psychologists enunciate the view that people enjoy distinguished cognitive privileges vis-à-vis their own affective states.

C. Jäger (✉)
Department of Philosophy, University of Innsbruck, Karl-Rahner-Platz 1,
6020 Innsbruck, Austria
e-mail: christoph.jaeger@uibk.ac.at

"Whether a man be asleep or dreaming, he could not feel sad or moved by some other passion, were it not very true that the soul has in it that passion."
(Descartes, *The Passions of the Soul*, 1649, art. 26, AT, p. 349)

"The objects of… [consciousness] are our present pains, our pleasures, our hopes, our fears, our desires, our doubts, or thoughts of every kind; in a word, all the passions, and all the actions and operations of our own minds, while they are present." "When a man is angry… he is conscious of it." "We believe what we are conscious of."
(Thomas Reid, *Essays on the Intellectual Powers of Man*, 1785, VI, Chap. 5, p. 470; I, Chap. 2: 'Principles taken for granted', p. 42; and VI, Chap. 5, p. 497)

"With respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurable surpassing those that can be possessed by any one else."
(John Stuart Mill, *On Liberty*, 1859, Chap. IV, p. 137)

"It is surely of the essence of an emotion that we should be aware of it, i.e. that it should become known to consciousness. Thus the possibility of the attribute of unconsciousness would be completely excluded as far as emotions, feelings, and affects are concerned."
(Sigmund Freud, 'The Unconscious', 1915b, p. 177)

"If I am angry, then I, but no-one else, knows that I am angry just by virtue of my being angry."
(William P. Alston, 'Emotion and Feeling', 1967, p. 482)

"For every *x*, if *x* has the property of being sad, and if *x* considers the question whether he is sad, then it is certain for *x* that he is sad."
(Roderick Chisholm, *The Foundations of Knowing*, 1982, p. 14)

"I would argue that it is not possible to have an unconscious emotion."
(Gerald L. Clore, 'Why Emotions are Never Unconscious', 1994, p. 285)

"In my view, 'emotions' are affectively charged, subjectively experienced states of awareness. Emotions, in other words, are conscious states. … Nonconscious emotions do not exist."
(Joseph E. LeDoux, 'Emotional Processing, But Not Emotions Can Occur Unconsciously', 1994, p. 291)

"[The concept of an unconscious emotion] becomes logically impossible because it provides a contradiction in terms."
(Arne Oehman, 'Distinguishing Unconscious from Conscious Mental Processes: Methodological Considerations and Theoretical Implications', 1999, p. 334)

"In the sense of consciousness as awareness, every emotion is necessarily conscious."

(Robert Solomon, 'Emotions, Thoughts, and Feelings: What is a "Cognitive Theory" of the Emotions, and Does it Neglect Affectivity?', 2003, p. 5)

"A careful consideration of the relevant evidence falls short of supporting the claim that there are unconscious emotions."
(Anthony Hatzimoysis, 'The Case Against Unconscious Emotions', 2007, p. 292)

These quotations range over a period of some 350 years. Their wording varies, but they concur in the view that, in one way or another, a person's cognitive access to his own emotions is superior to his access to other people's emotions.[1] A common rationale for this claim is that third-person attributions of emotional states must rely on inference and empirical evidence, whereas corresponding first-person attributions need not and usually do not rest on any such basis. Hence the latter would appear to be immune to a variety of errors which often befall corresponding third-person ascriptions.[2] However, this view, although *prima facie* plausible and attractive, needs a reality test. In what follows I argue that it doesn't pass. The truth is that we often suffer from affective ignorance and fail to have the kind of privileged access to our emotional lives that philosophers and psychologists, both from rationalist and empiricist camps, have traditionally maintained we have. In order to see why this is so, we shall need to have before us a more detailed picture of the alleged cognitive privileges. First, however, some items of business.

The above quotations talk about "passions", "feelings", and "emotions". Another expression that frequently figures in discussions of our subject is "affective state". I will use this latter locution as an umbrella term that covers affective phenomena such as "basic" emotions (fear, anger, joy, sadness, disgust, or shame); "hot" varieties of them (or of other emotions, for example boiling with rage or feeling giddy with elation); "cold" emotions (feeling mild forms of envy, admiration); cognitively complex emotions or emotion-like states (indignation, resentment, patriotism); and also moods or mood-like states (feeling melancholic, cheerful, etc.).[3]

[1] The quotation from Alston represents what he regards as a *prima facie* view that he does not fully endorse. It may also be noted that Freud's remarks on the topic are notoriously incoherent: Immediately after the passage quoted above he says that "in psychoanalytic practice we are accustomed to speak of unconscious love, hate, anger, etc., and find that it is impossible to avoid the strange conjunction, 'unconscious consciousness of guilt', or paradoxical 'unconscious anxiety'" (Freud 1915b, p. 177f).

[2] Not every philosopher endorses this claim. Recent sceptics include Daniel Haybron (2008) and Jordi Fernández (2003). Haybron, from whom I have borrowed the title-term of this paper (p. 200, and *passim*), argues that often we don't know how happy we are. Although he opts for an anti-hedonist view of happiness, he does regard happiness to be an affective state. I will discuss some of Haybron's arguments in Sect. 3. Fernández remarks in passing that, "arguably, we do not know in any special way about our character traits and *some* of our emotions" (2003, p. 353, my emphasis).

[3] Some of these "states" may perhaps better be characterized as mental episodes. (Note that Reid talks about "actions and operations" of the mind.) For ease of exposition I shall often just say "states". Wherever "episode" or "process" appears more appropriate to the reader she can accordingly rephrase my arguments without affecting their central point.

Affective predicates can be used to ascribe either occurrent mental states (or episodes) or mental dispositions. Very long-term forms of the latter may culminate in character traits. It may be argued from the outset that dispositions are unsuitable candidates for privileged access. However, the above quotations are best interpreted as referring to occurrent emotions. Since the latter would in any event appear to be the best candidates for privileged access I shall restrict my discussion to affective states in the occurrent-mental-state sense.

Note that some, but not all, authors from the above list explicitly employ epistemic notions. Reid, Mill, Freud, Alston, and Chisholm talk about people knowing, or feeling certain about, their emotions. Others maintain that emotions cannot remain "unconscious" (Clore, Oehman, Hatzimoysis), or that they must be conscious in the sense of "awareness" (Freud, Solomon, LeDoux). I propose to construe all these statements in terms of knowledge and belief. For example, when Chisholm talks about "certainty" what is at issue is subjective certainty, or maximally firm belief. I shall henceforth omit this qualification and use "belief" exclusively to denote firm belief.

Beliefs need not be conscious in the sense in which occurrent phenomenal states are conscious. The reverse, however, does hold. As Reid points out, "we believe what we are conscious of". Reid also argues for the stronger claim that "by consciousness we *know* certainly the existence of our present thoughts and passions" (Essays, p. 42, emphasis C.J.). I suggest that we follow Reid and, in addition to notions such as "certainty", also construe claims in which emotions are said to be necessarily conscious in terms of knowledge and belief. Similarly for awareness. *S*'s being aware of *X* is naturally construed as *S*'s being in a mental state which entails that *S* has knowledge about *X*.[4]

2 Emotions and Special Access

A generalized version of Chisholm's claim about sadness may be put by saying that, whenever a person is in some emotional state (of an appropriate kind), she strongly believes that she is in that state. Stated this way, however, the principle only captures a version of what we may call "positive affective introspection". Such claims are often supplemented with corresponding statements about "negative introspection". In the present case, the corresponding negative thesis is that, if *S* is not in a given emotional state, *S* believes herself not to be in that state. Let us call the conjunction of these two principles (Affective) Self-Intimation.

Chisholm, along with other authors from the above list, does not explicitly formulate his view as a conceptual claim. Yet this seems to be what he intends. In general we may follow Clore, Oehman, Solomon, and others, and construe Self-Intimation as the modal thesis that it is (logically) impossible for the relation in question not to obtain. The first principle to be considered, then, is this:

[4] This is in line with what Alston (1989) proposes in his classic discussion of varieties of privileged access.

(Affective) Self-Intimation:

For each affective state *A* (of an appropriate type) it holds that, necessarily,
(i) if *S* is in *A*, *S* believes herself to be in *A*; and
(ii) if *S* is not in *A*, *S* believes herself not to be in *A*.

I have omitted time indices, but let it tacitly be read that *S*'s beliefs occur at the same time at which *S* is (is not) in *A*. Moreover, as it stands, Self-Intimation does not express any first-person privilege. The alleged privilege is that only *S* him- or herself, and no one else, is related in this way to *S*'s affective states. For the sake of brevity I shall from now on omit this addendum but assume that the idea is that only the subjects of the states in question bear the respective relations to these states.

Another influential variety of special access "reverses the direction" of Self-Intimation. The passage from Descartes, for example, can be interpreted as endorsing a view that entails

(Affective) Infallibility:

For each affective state *A* (of an appropriate type) it holds that, necessarily,
(i) if *S* believes herself to be in *A*, *S* is in *A*; and
(ii) if *S* believes herself not to be in *A*, *S* is not in *A*.

Often Self-Intimation and Infallibility are treated as if they were logically independent of each other. It is worth noting, however, that the latter principle is entailed by the former, provided we accept what may be called the "law of doxastic non-contradiction":

Doxastic Non-Contradiction (=DNC):

Necessarily: If *S* believes not-*p*, then *S* does not (at the same time) believe *p*;
or, equivalently:
Necessarily: If *S* believes *p*, then *S* does not (at the same time) believe not-*p*.

Given DNC, Self-Intimation, part (ii), entails Infallibility, part (i), and Self-Intimation (i) entails Infallibility (ii).[5] I realize that, read as a general empirical claim about actual human cognizers, DNC is most certainly false. Yet the principle may be regarded as a requirement for minimally (epistemically) rational subjects. Hence at least under this restriction, Infallibility is just a corollary of Self-Intimation. I will return to DNC in Sect. 4.

The conjunction of Self-Intimation and Infallibility is equivalent to

[5] Given DNC, the consequent of Self-Intimation (ii) entails that *S* does not believe that *S* is in *A*. Thus, by transitivity of entailment, Self-Intimation (ii) entails Self-Intimation (ii)*: If *S* is not in *A*, *S* does not believe herself to be in *A*, which is equivalent to Infallibility (i). Similarly, if DNC holds, the consequent of Self-Intimation (i) entails that *S* does not believe herself not to be in *A*. Hence Self-Intimation (i) entails Self-Intimation (i)*: If *S* is in *A*, then *S* does not believe herself not to be in *A*, which is equivalent to Infallibility (ii).

(Affective) Omniscience:

For each affective state A (of an appropriate type) it holds that, necessarily,
(i) S is in A iff S believes so; and
(ii) S is not in A iff S believes she is not.

Since Self-Intimation and DNC entail Infallibility, they also entail Omniscience.

Some statements from the above list talk about people's *knowledge* of their own emotions (Mill, Freud, Alston). If we complement again the "positive" version of this idea with a corresponding negative clause, we obtain the principle of

Perfect (Affective) Knowledge:

For each affective state A (of an appropriate type) it holds that, necessarily,
(i) if S is in A, S knows that she is; and
(iI) if S is not in A, S knows that she is not in A.

Since knowledge entails belief, Perfect Knowledge entails Self-Intimation. The same holds for principles that, instead of focussing on knowledge, maintain special-access claims in terms of justified (or warranted, rational, epistemically apt) belief. Such principles, too, entail Self-Intimation and thus, like Perfect Knowledge, together with DNC entail Omniscience. Table 1 summarizes the logical relations considered so far.

Having a clear idea about the logical relations between various privileged access claims may reduce the complexity of investigations about whether or not such claims are *true*. Suppose, for example, it could be shown that Infallibility is false. Then Omniscience and, for subjects who conform to DNC, Self-Intimation, Perfect Knowledge, and other principles that are logically stronger than Self-Intimation would go by the board as well. I conclude this section with one more observation of this kind.

For each affective state A and subject S it holds that S either is or is not in A–$A(S)$ or $\sim A(S)$, *tertium non datur*. Together with Self-Intimation this directly entails what may be called:

Table 1 Logical relations among prominent doxastic and privileged access principles

Perfect Knowledge	$\Rightarrow$	Self-Intimation				
		Self-Intimation & DNC	$\Rightarrow$	Infallibility		
				Infallibility & Self-Intimation	$\Leftrightarrow$	Omniscience
		Self-Intimation & DNC	$\Rightarrow$			Omniscience

Doxastic Excluded Middle (for affective states) (=DEM):

For each affective state A (of an appropriate type) it holds that, necessarily, S believes herself to be in A, or S believes herself not be in A.

Borrowing an expression from Sydney Shoemaker (1990, p. 51f), we may say that this principle puts a "ban on first-person agnosticism" regarding affective states. It tells us that belief suspension is impossible in such cases. However, $[A(S)$ or $\sim A(S)]$ is just an instantiation of bivalence. Should it turn out therefore that DEM is false, either Self-Intimation or bivalence would have to be jettisoned. It is bivalence, I dare say, which would in that case be in the stronger position.[6]

I now turn to the question of whether any of these principles are true, at least when applied to paradigmatic affective states.

3 Affective Ignorance and Third-Person Authority

On closer inspection both Infallibility and DEM—and thus also Perfect Knowledge, Self-Intimation, and Omniscience—turn out to be false. The truth is that we often seriously misinterpret, misrepresent, repress, and ignore our own affective states. Let it be emphasized from the outset that my thesis differs from the following related but weaker claim. Peter Goldie has recently argued that it is possible to "have an emotion" and yet not "feel" it in the sense of being "reflectively conscious" of it (2000, pp. 62–67). I agree, but I will argue for a stronger tenet. Goldie elucidates his use of "reflective consciousness" by way of examples: "being aware that I feel afraid; being aware that I feel afraid of the oncoming vehicle; and being aware that I am thinking of the oncoming vehicle as out of control…" (2000, p. 64). He then argues that one can be so absorbed in a situation that, even though one is emotionally engaged in it, one lacks reflective consciousness of this fact. But Goldie is here using the notion of being reflectively conscious of X in the emphatic sense of the subject's currently thinking, or being attentively aware, of X. However, taken in this sense, believing that p does not entail being reflectively conscious of the fact that p. Hence lacking reflective consciousness of an emotional state one is in does not rule out that one believes oneself to be in that state. I want to argue for the falsehood of even those special access views according to which having an emotion entails *believing* so, or according to which believing that one has a certain emotion entails that this belief is true. I shall now present five arguments for this claim: Belief Suspension; Interpersonal Variance; Adaptation; Alexithymia; and Repression. Some of these arguments will be formulated on a more or less intuitive level; others will draw on what seems to be ample empirical evidence.

[6] I will not comb through all the logical relations that may be considered in this context. Further interesting relations are: DEM $\Rightarrow$ (Infallibility $\Rightarrow$ Self-Intimation); DEM $\Rightarrow$ (Infallibility $\Leftrightarrow$ Omniscience); Infallibility $\Rightarrow$ (DEM $\Leftrightarrow$ Omniscience). Let DNC* be: $\mathrm{Bel}(S, A(S)) \Rightarrow \sim\mathrm{Bel}(S, \sim A(S))$. We then also have: Infallibility $\Rightarrow$ DNC* and DEM & DNC* $\Rightarrow$ (Self-Intimation $\Leftrightarrow$ Infallibility $\Leftrightarrow$ Omniscience). Let's save the proofs for a rainy afternoon.

3.1 Belief Suspension

To begin with, it can hardly be denied that we often do suspend belief about our current emotions and other affective states and processes. We frequently shrug our shoulders when asked whether we are happy (anxious, jealous, etc.). It may be conceded that sometimes this will be due to the desire to meet social expectations. But I can see no reason always to take people to be insincere when they express agnosticism as to whether or not they are currently in a certain emotional state. If that is right, then DEM (and hence Self-Intimation) is false.

Note that the conclusion that DEM is false is indirectly supported by so-called first-person authority theses. These are weaker than the strong "Cartesian versions" of privileged access. Authority claims merely maintain that, even though misinterpretation of one's current mental states is possible, this is the exception. Applied to emotions, the idea is that we are generally justified in taking people's reports about their current emotions to be veridical. Donald Davidson (1987) has famously argued for such a claim regarding thoughts (and other mental states). "Our sincere avowals concerning our present states of mind", he writes, "are not subject to the failings of conclusions based on evidence. Thus sincere first person present-tense claims about thoughts, while neither infallible nor incorrigible, have an authority no second or third person claim, or first person other-tense claim, can have" (p. 441). However, a person's sincere avowal that she is (or is not) in a mental state *M* expresses her belief that she is (is not) in *M*. Accordingly, her sincere report that she is *unable* to tell whether or not she is in *M* informs the hearer that she withholds belief as to whether or not she is in *M*. Now if, as Davidson and many other authors have argued, there are good reasons for regarding first-person reports expressing people's beliefs or disbeliefs about their own minds as authoritative, then there would seem to be equally good reasons for regarding reports as authoritative in which people claim that they suspend belief as to whether or not they are in *M*. In short, if first-person authority claims are acceptable, then we are also entitled to regard announcements as veridical in which people report *not* to know whether or not they are having a certain emotion. If so, we are entitled to believe that DEM is false.

In spite of being more moderate than traditional privileged access principles, first-person authority claims are friendly amendments of Cartesian accounts of the mental. I shall now argue that, on closer inspection, even mere authority theses fail when applied to emotions. Here is a story once told by William Lyons:

> "The following seems to be an uncontroversial case of emotions: O'Reilly is so taken up by the discussion at the curriculum meeting that he does not realize that he is becoming very angry with Macdonald who is suggesting that the central texts in the first year course should consist only of the writings of the Existentialists. It is only later on, when Macdonald curtly remarks to O'Reilly that there was no need to get so heated, and he overhears MacFee wonder why he got so angry, that O'Reilly realizes that he must have become very angry during the meeting." (Lyons 1980, p. 6)

I think that Lyons is right in maintaining that such situations frequently occur (indeed especially in department meetings). He argues that if a feeling theory of the emotions in a Cartesian mould were true, the above case would be impossible. Feelings, Lyons says, are subjective states in which one is aware of whatever is their content (ibid.). The crucial point for our present question is that O'Reilly himself, initially unaware of his emotion, accepts *third-person reports* about his affective state as authoritative. Prior to being confronted with Macdonald's and MacFee's remarks, O'Reilly would not have conceded that he had become cross with Macdonald. Moreover, it seems that O'Reilly is perfectly rational and justified (warranted, epistemically in the clear) when he accepts Macdonald's and MacFee's diagnosis. The moral is that third-person assessments of an individual's affective states can—and often should—be taken to be more authoritative than first-person beliefs and (potential) first-person reports about such states.

3.2 Interpersonal Variance

Next consider the fact that there is remarkable interpersonal variance in people's perceptiveness vis-à-vis their moods and emotions. This has recently been stressed by Daniel Haybron. We are not all Prousts. As Haybron illustrates, "surely the Dalai Lama notices things about his emotional state that would have completely escaped John Wayne" (Haybron 2008, p. 206). For example, certain moods and affective background states tend to be elusive for those of us who do not or cannot care about these states. A good example is unnoticed anxiety or unconscious stress. As Haybron points out, stress, if it occurs in certain qualities and quantities, tends to spoil one's happiness. But we are often unaware that this happens. There may be an issue here about whether happiness is an emotion. The point remains, however, that anxiety and related forms of negative affect may go unnoticed *even when the subject focuses her attention* on the question whether or not she is in such states. If that is true, we do not even generally enjoy "access consciousness" to such states. People engage in all sorts of psychological training and therapy precisely because they wish to acquire more sensitivity to affective states and dispositions which they believe may have negative effects on their well being. The underlying assumption clearly is that such states fail to be suitable objects of the kind of privileged access captured, for example, in Infallibility (ii). But if Infallibility turns out to be false, then Self-Intimation, Perfect Knowledge, and Omniscience must go as well. The following arguments will support this claim in more detail. I shall look at three empirical studies.

3.3 Affective Adaptation

First, let us consider an issue again raised in Haybron (2008, pp. 205 f). Empirical evidence suggests that people tend to adapt themselves to stress and anxiety, at least if such conditions persist over relatively long periods of time. Studies have been conducted in which individuals worked for three hours in a simulated office. One group was subjected to low-intensity noise, the control group was not (Evans and Johnson 2000). Physiological and behavioral stress indicators were found to

be significantly higher in noisy-office subjects than in those who worked in a quiet office. However, the self-reports that asked people whether they felt "bothered, worried, relaxed, frustrated, unhappy, contended, [or] tense" did not vary between the two conditions. This suggests that the noisy-office subjects failed correctly to perceive their levels of affective arousal. Interestingly, in an earlier study, when less experienced office workers had been exposed to noise, yet for a much shorter time, the levels of perceived stress did vary between the noisy-office and the quiet-office group. A plausible explanation is that the experienced workers in the later study had adapted to persistent negative affect. Put in terms of our privileged access principles laid out in Sect. 2, the moral is again that these findings suggest that Infallibility (ii), and hence also Self-Intimation, Omniscience, and Perfect Knowledge, are false.

3.4 Alexithymia

Next, consider alexithymia. Traditional characterizations of alexithymia tended to echo the literal meaning of the term: "absence of words for emotion". Accordingly, the phenomenon was originally defined as a deficit in the capacity to verbally describe and identify feelings or emotions.[7] Various studies carried out since the mid 1990s, however, strongly suggest that alexithymia does not just involve impairment at the linguistic level. According to what is now a widely accepted characterization, alexithymia has four salient features: "(1) difficulty identifying feelings and distinguishing between feelings and the bodily sensations of emotional arousal; (2) difficulty describing feelings to other people; (3) constricted imaginal processes, as evidenced by a paucity of fantasies; and (4) a stimulus bound, externally orientated cognitive style" (Taylor et al. 1997, p. 29; see also Taylor 2004). It is evident from the authors' explicit distinction between feelings and sensations that in this characterization "feeling" doesn't refer merely to bodily sensations but to what in this paper has been referred to as "emotion".

Features (3) and (4) may not have any direct bearing on our topic. But, to begin with, (2) does not blend well with first-person authority views. If I cannot properly describe my emotions to other people, how can my reports deserve to be taken with special authority? Moreover, a natural interpretation of the inability to "identify" one's emotional state is that, when one is in such a state, one is unable to form a correct belief about this. Hence feature (1) of the above characterization is inconsistent with Self-Intimation.

The hypothesis that alexithymia is an impairment that reaches beyond the linguistic level was empirically confirmed. Recent studies have been summarized by saying that "the hallmark of alexithymia, a difficulty in putting emotion into words, may be a marker of a more general impairment in the capacity for emotion information processing" and "the capacity to recognize emotions" (Lane et al. 1996, p. 203). In a community survey 380 individuals completed two measures of alexithymia, the so-called Levels of Emotional Awareness Scale (LEAS) and the Toronto Alexithymia Scale (TAS-20). In addition, the subjects completed the

[7] See Peter E. Sifneos's characterization in his pioneering paper from 1973 (p. 255).

Perception of Affect Task (PAT). The sample was balanced across gender, three socio-economic groups, and five age groups. TAS-20 is a self-report measure widely used to test individuals for alexithymia. It contains items such as "I find it hard to describe how I feel about people", "I often don't know why I am angry", etc. If subjects score ≥61 on a scale for a maximum score of 100, they are classified as alexithymic (intermediate 52–60, nonalexithymic ≤51).

LEAS asks subjects to report their anticipated emotions and those of another person in 20 different scenes. For example, one story to be commented on is that "you and your best friend are in the same line of work. There is a prize given annually to the best performance of the year. The two of you work hard to win the prize. One night the winner is announced: your friend. How would you feel? How would your friend feel?" Responses for each scene receive a score of 0–5, the criteria being whether, and in which way, words like "feel" are employed. Finally, PAT is a questionnaire that asks subjects to identify emotions such as happiness, sadness, fear, anger, disgust, etc., which are circumscribed in short stories, or presented in photographs of human faces or of "emotional scenes" where no human faces are visible.[8]

It was found that both the TAS-20 and LEAS correlated substantially with PAT. More specifically, the major findings have been summarized by saying that "the ability to recognize emotions decreases as alexithymia scores increase and… this decreased ability is both verbal and nonverbal" (Lane et al. 1996, p. 207). If this interpretation is correct, these studies provide direct evidence for the view that alexithymic individuals fail to have adequate beliefs about their emotions.

(v) Repression. Finally, consider emotional repression. This phenomenon may be closely related to alexithymia, but in the psychological literature it is usually treated as a distinct affair. The notion was introduced by Freud, for whom wishes and desires are paradigm objects of repression. His official characterizations, however, tended to be fairly liberal,[9] and in general psychoanalysis does have room for repressed and unconscious emotions. Here I want to discuss an empirically operationalized concept of emotional repression. In recent years it could be shown that, as researchers have concluded, people often genuinely "deceive themselves about the nature of their emotional responses" (Weinberger and Davidson 1994, p. 589; cf. also Weinberger et al. 1979, and Weinberger 1990).

In a typical study, researchers combined two kinds of self-report measures to pre-classify subjects as repressors. People filled out a trait anxiety questionnaire (such as the Taylor Manifest Anxiety Scale), and the Marlowe-Crowne Social Desirability (=MC) scale. The former asks about perceived tendencies to experience anxiety and

[8] Responses are—depending on the subtask—to be chosen from a list of seven words that denote emotions, or from the two types of pictures just described. Altogether there are four forms of stimulus-response combinations: sentences-words (subtask 1: emotion words are to be assigned to sentences/stories); faces-words (subtask 2: emotion words are to be assigned to photographs of human faces); sentences-faces (subtask 3: photographs of human faces are to be assigned to the sentences/stories that also serve as stimuli in subtask 1); and faces-scenes (subtask 4: photographs from "emotional" scenes that do not depict human faces are to be assigned to photographs of human faces).

[9] In his famous paper "Repression", for example, Freud writes that "the essence of repression lies simply in turning something away, and keeping it at a distance, from the conscious" (Freud 1915a, p. 147).

anxiety-like emotions. The MC scale, by contrast, measures people's self-perception regarding behavior that is socially sanctioned or approved. (This scale contains items such as "I never hesitate to go out of my way to help someone in trouble".) Interestingly, it could be shown that individuals who scored high on the MC scale did not twist answers in order to meet social expectations, but actually believed themselves to conform to the unrealistically high standards they reported to meet. It was plausibly concluded that scoring high on the MC scale indicates a high level of defensiveness. Moreover, it was hypothesized that people who scored high on the MC scale but low on the anxiety scale may in fact not be as invulnerable to emotional disturbances as they reported. Since their high MC scores indicated a high level of defensiveness, it seemed likely that their self-estimation regarding the tendency to experience anxiety did not reflect the true level of proneness to emotional distress.

Experiments confirmed this hypothesis. In a second round the subjects were confronted with affectively charged phrases with aggressive or sexual content, with footage of industrial accidents, or with modified forms of the prisoner's dilemma.

Table 2 Experimental design of a typical repression study

Experimental results

Marlowe-Crowne-Scale of defensiveness	Anxiety scale (self-report): low	Anxiety scale (self-report): high
high	Repressors high levels of physiological arousal, low levels of self-reported emotion	Defensive high anxious few subjects (often discarded from statistical analysis), intermediate results
low	Truly low anxious low levels both of physiological arousal and self-reported emotion	High anxious high levels both of physiological arousal and self-reported emotion

Physiological reactions such as changes in heart rate, blood pressure, skin conductance, etc., and behavioral reactions such as changes in facial expressions, speech patterns, reaction times, or objective test performances were measured. In addition, subjective perception of affective reactions was again assessed by self-report measures. The result was that those individuals who had been preselected as repressors showed physiological and behavioral reactions to stressors that were as intense as, or even more intense than, the reactions of the control groups. However, subjects from the repressor group reported no, or only a very low level of, emotional distress when confronted with the stimuli. Table 2 summarizes the setup and the results of these studies.

A first series of experiments tested negative emotions, but further studies also included positive emotions. It was found that repressors report higher levels of affect than control subjects when they (the repressors) perceive no potential threat to their self-esteem, but that self-reported affect dramatically diminishes in ego-threatening situations. Hence response dissociation also occurs intrapersonally when the situation variable is manipulated (Mendolia 1999). In summary, it was concluded that dissociation between objective and subjective reactions to emotion-eliciting situations suggests the existence of a repressive coping style that tends to be adopted by people who are hypersensitive to both negative and positive emotional stimuli. What are the consequences for affective special access?

Again, these findings suggest that Infallibility and, so the refrain goes, thus also Self-Intimation, Perfect Knowledge, and Omniscience, are false. Repressors tend to believe that they are not anxious (exulted, sad, etc.) even when they are in these states, and they are often anxious (exulted, sad, etc.) without noticing this.

4 Objections and Conclusion

I conclude by looking at five objections concerning the foregoing arguments. (i) One general worry may be that the studies I have discussed all rely on self-report measures. So don't they invoke special access assumptions after all? The questionnaires measure what subjects believe about their emotional lives. But this is consistent with the view that such beliefs are often false. Moreover, the studies don't demonstrate that we are all repressors; that we all suffer from alexithymia; have adapted to persistent negative affect, and so on. But even though for practical purposes certain cut-off points had to be defined, on a theoretical level concepts such as "alexithymia", "repression", and "adaptation" are generally regarded as referring to dynamic psychological properties. This seems very plausible. While presumably not everyone on every occasion completely ignores or misinterprets all of his or her emotions, many of us get it a bit wrong sometimes.

(ii) In Sect. 2 I conceded that DNC, the principle of doxastic non-contradiction, should not be interpreted as an unrestricted empirical claim ranging over all (normal) human adults. Yet I employed DNC to derive Infallibility from Self-Intimation and then argued that there is strong empirical evidence against Infallibility (and hence, if we accept DNC, also against Self-Intimation and Perfect

Knowledge). Was this move dialectically legitimate? Yes it was, but a clarifying comment may be in order. The point is indeed only that individuals who conform to Affective Self-Intimation *and who meet DNC* also conform to Affective Infallibility. But that is coherent with my overall argument. If only individuals who, in the relevant epistemic contexts, did *not* conform to DNC were amenable to affective adaptation, inaccurate affective awareness, etc., then it could not legitimately be concluded from the studies that, since those individuals violate Affective Infallibility, they also violate Affective Self-Intimation. However, it would be highly implausible to assume that only those of us who don't meet the rationality requirement of not simultaneously believing both *p* and not-*p* adapt themselves to stress and anxiety, are liable to repression, and so on.

(iii) Anthony Hatzimoysis (2007) has recently argued that alexithymia in fact does not jeopardize the view that emotions are necessarily conscious. To rebut the thesis that alexithymic individuals are conscious of their emotions, he says, one would have to show that "the subject is not aware of her experience unless, first, she knows what caused the experience; and secondly, she is able to conceptualize and express in linguistic medium the nature of that experience" (p. 295). Yet, the argument goes, such a view would misconstrue what it means to be aware of an experience.

I don't wish to dispute that being aware of an *experience* presupposes neither having knowledge about its cause nor being able properly to express its nature in linguistic medium. But this is not what is at issue. The question is whether alexithymic individuals are aware of their *emotions*. Empirical research on alexithymia has not established that alexithymic individuals are unaware of any kind of affective experience while undergoing an emotion. Yet what they lack is the ability properly to assign certain experiences to emotional states. Hatzimoysis's argument would only go through if he adopted a theory according to which having an emotion reduces to having certain affective experiences. But, first, he himself does not seem to endorse such a strong form of anti-cognitivism. Second, pure feeling theories face a host of well known systematic objections. I will not rehearse the many arguments that have been urged against such theories in recent decades. But we can draw at least the following conditional conclusion: If we don't want to adopt a pure feeling theory, then Hatzimoysis's argument, while it may be on target with respect to affect-related experiences, fails as an objection to the view that alexithymic individuals lack privileged access to their emotions.

(iv) A *prima facie* plausible suspicion may nevertheless be that the natural habitats of affective special access claims are feeling-based accounts of the emotions. These identify emotions with certain phenomenal experiences, and the latter are certainly among the most promising candidates for special access. As already indicated, however, feeling theories have increasingly fallen into disfavor,[10] and there has been a strong shift towards cognitivist and judgementalist approaches in recent decades. With the decline of feeling theories, have not affective privileged

[10] Contrary tendencies, however, can be found in two careful recent approaches by Jesse Prinz (see his "embodied appraisals" theory, sketched in 2003 and 2004) and Jenefer Robinson (2004), who argues that emotions are "affective appraisals" of bodily feelings.

access views lost much of their former attraction as well, and doesn't this paper thus launch a foray into abandoned territory?

Not at all. Affective special access views are not confined to feeling theories. They seem to enjoy just as much popularity among cognitivists. For example, in Sect. 1 I quoted the late Robert Solomon who, despite his special access thesis, also advocated a radical form of cognitivism about emotions. "An emotion", Solomon holds, "*is* an evaluative (or 'normative') judgment" (1993, p. 126, emphasis C.J.). So, theories of this stripe hold that we enjoy special cognitive access to (certain kinds of) judgements. A closely related view has also widely been defended in recent decades with respect to beliefs and other propositional attitudes. In the 1980s content externalism was accused of being incompatible with privileged self-knowledge. Most philosophers who worked on the topic, however, were not prepared to give up special access and authority views regarding thoughts and beliefs. Instead, they sought to reconcile such views with externalism. Today a widely accepted stance on the issue is compatibilism (about privileged self-knowledge and externalism). If this position is well motivated, there should be no reason to suppose that cognitive theories of the emotions as such are any less hospitable to affective special access theses than traditional feeling theories.

(v) Finally, a potential worry may be that the phenomena investigated in the studies may not properly be called "emotions" or "affective states" in the first place. It may be argued that emotions are *by definition* self-intimating, that they constitute objects of infallible access, and so on. In that case any argument that seeks to demonstrate that this is not so must be talking about something else. Such a view seems to be suggested, for example, in the above quotations from Freud ("it is of the essence of an emotion that we should be aware of it"); Clore ("it is not possible to have an unconscious emotion"); Oehman ("[the concept of an unconscious emotion] provides a contradiction in terms"), and Solomon ("every emotion is necessarily conscious").

It may of course be stipulated that terms like "emotion" and "affective state" be used in that way. But what we are after is not some stipulative notion but the phenomena we normally call by the name "emotion". If that is what Freud, Clore, Oehman, Solomon, and friends have in mind, however, they owe us an alternative description of the cases discussed in this paper. Consider repression. Suppose we refuse to countenance the term "emotion" for the repressed states investigated in the studies. What, then, is their subject? The critic must redescribe the phenomena in non-emotional terms. But it seems doubtful that this could be done in a plausible way. Note that, first, it would have to be conceded that from a third-person perspective the mental states in question are indistinguishable—except for differences in the subject's verbal behavior—from emotional states that are not being repressed. Repressed emotional (or quasi-emotional) states have all the physiological and behavioral characteristics that conscious emotions have. Secondly, the subjects of the truly high anxious control group, who displayed similar patterns of behavior and similar levels of physiological arousal than the repressors, did classify their states as states of emotional arousal. For them, there was no question that they had become anxious and excited in the test situations. Considering these facts, it seems fair to say that the burden of proof is with the

objector, who refuses to apply the term "emotion" to the states in question. Until we see a convincing argument for such a stance, however, I think we have good reason to conclude that emotions often act behind the stage of Cartesian Theatre. The heart—sometimes—has reasons that reason doesn't know.

Acknowledgments For helpful comments I am indepted to Dan Haybron, Nick Shackel, Christian Tapp, and two anonymous referees for *Erkenntnis*.

References

Alston, W. P. (1967). Emotion and feeling. In P. Edwards (Ed.), *The encyclopedia of philosophy* (Vol. 2, pp. 479–486). New York & London: Macmillan Company.

Alston, W. P. (1989). Varieties of privileged access. In W. P. Alston (Ed.), *Epistemic justification: Essays in the theory of knowledge*. Ithaca & London: Cornell University Press.

Chisholm, R. M. (1982). *The foundations of knowing*. Minneapolis: University of Minnesota Press.

Clore, G. L. (1994). Why emotions are never unconscious. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 285–290). New York: Oxford University Press.

Davidson, D. (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, *60*, 441–458 (Reprinted in Q. Cassam (Ed.), *Self-knowledge*. Oxford: Oxford University Press).

Descartes, R. (1989). *The passions of the soul*. (S. H. Voss, Trans.) Indianapolis: Hacket.

Evans, G. W., & Johnson, D. (2000). Stress and open office noise. *Journal of Applied Psychology, 85*, 779–783.

Fernández, J. (2003). Privileged access naturalized. *The Philosophical Quarterly, 53*, 352–372.

Freud, S. (1915a). Repression. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 14, pp. 143–158). London: The Hogarth Press.

Freud, S. (1915b). The unconscious. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 14, pp. 161–215). London: The Hogarth Press.

Goldie, P. (2000). *The emotions—a philosophical exploration*. Oxford: Clarendon Press.

Hatzimoysis, A. (2007). The case against unconscious emotions. *Analysis, 67*, 292–299.

Haybron, D. M. (2008). *The pursuit of unhappiness—the elusive psychology of well-being*. Oxford: Oxford University Press.

Lane, R. D., et al. (1996). Impaired verbal and nonverbal emotion recognition in alexithymia. *Psychosomatic Medicine, 58*, 203–210.

LeDoux, J. (1994). Emotional processing, but not emotions can occur unconsciously. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 291–292). New York: Oxford University Press.

Lyons, J. (1980). *Emotion*. Cambridge: Cambridge University Press.

Mendolia, M. (1999). Repressors' appraisals of emotional stimuli in threatening and nonthreatening positive emotional contexts. *Journal of Research in Personality, 33*, 1–26.

Mill, J. S. (1859). *On liberty*. London: John W. Parker & Son.

Oehmann, A. (1999). Distinguishing unconscious from conscious emotional processes: Methodological considerations and theoretical implications. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 321–351). Chichester: John Wiley & Sons Ltd.

Prinz, J. (2003). Emotion, psychosemantics, and embodied appraisals. In A. Hatzimoysis (Ed.), *Philosophy and the emotions* (pp. 69–86). Cambridge: Cambridge University Press.

Prinz, J. (2004). Embodied emotions. In R. C. Solomon (Ed.), *Thinking about feeling: Contemporary philosophers on emotions* (pp. 44–58). Oxford & New York: Oxford University Press.

Reid, T. (2002). *Essays on the intellectual powers of man*. Edinburgh: Edinburgh University Press.

Robinson, J. (2004). Emotion: Biological fact or social construction? In R. C. Solomon (Ed.), *Thinking about feeling: Contemporary philosophers on emotions* (pp. 28–43). Oxford & New York: Oxford University Press.

Shoemaker, S. (1990). First-person access. (Reprinted in Shoemaker, S., *The first-person perspective and other essays* (pp. 50–73). Cambridge: Cambridge University Press).

Sifneos, P. E. (1973). The prevalence of "Alexithymic" characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics, 22*, 255–266.

Solomon, R. C. (1993). *The passions*. Indianapolis: Hackett.

Solomon, R. C. (2003). Emotions, thoughts, and feelings: What is a "cognitive theory" of the emotions and does it neglect affectivity? In A. Hatzimoysis (Ed.), *Philosophy and the emotions* (pp. 1–18). Cambridge: Cambridge University Press.

Taylor, G. J. (2004). Alexithymia: Twenty-five years of theory and research. In I. Nyklicek, L. Temoshok, & A. Vingerhoets (Eds.), *Emotional expression and health: Advances in theory, assessment, and clinical applications* (pp. 137–138). Andover, UK: Brunner-Routledge.

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (1997). *Disorders of affect regulation: Alexithymia in medical and psychiatric illness*. Cambridge: Cambridge University Press.

Weinberger, D. A. (1990). The construct validity of the repressive coping style. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health*. Chicago & London: The University of Chicago Press.

Weinberger, D. A., & Davidson, M. N. (1994). Styles of inhibiting emotional expression: Distinguishing repressive coping from impression management. *Journal of Personality, 62*, 587–613.

Weinberger, D. A., Schwartz, G. E., & Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology, 88*, 369–380.