META-EMOTIONS

Christoph JÄGER & Anne BARTSCH
University of Aberdeen & University of Halle

Summary
This paper explores the phenomenon of meta-emotions. Meta-emotions are emotions people have about their own emotions. We analyze the intentional structure of meta-emotions and show how psychological findings support our account. Acknowledgement of meta-emotions can elucidate a number of important issues in the philosophy of mind and, more specifically, the philosophy and psychology of emotions. Among them are (allegedly) ambivalent or paradoxical emotions, emotional communication, emotional self-regulation, privileged access failure for repressed emotions, and survivor guilt.

1. Here's looking at you, kid!

Emotions, it is often claimed, are either positive or negative. According to a popular view both in philosophical and psychological research on emotions, joy, pleasure, and pride, for example, stand in stark contrast in this respect to fear, anger, sorrow, disgust, shame, and guilt. ¹ This division has a long and distinguished career. Aristotle characterizes the passions (πάθη) as affections of the soul that are accompanied by pain or by pleasure. “Such are anger, pity, fear,” he contends, “and all similar emotions and their contraries” (Rhetoric 1378a). Related ideas reverberate in current appraisal theories of the emotions, which postulate a close link

¹ See for example Gordon (1987), 27: “With a few exceptions, emotions are intuitively characterizable as ‘negative’ emotions, such as fear, embarrassment, and anger, or as ‘positive’ emotions, such as pride and gladness”. Similarly, William Lyons says in an influential study (1980, 91): “All emotions seem to include an evaluation, and an operative evaluation has to be either positive or negative”. Cf. also Green (1992), 171f., and statements in the same vein in Lazarus (1991) and Roseman (2001). See Solomon (2003) for a critical discussion of the positive and negative categories.
between the formation of emotions and appraisals of events and situations vis-à-vis their significance for one’s personal ends and goals. Situations that we think promote our goals, it is argued, usually elicit positive emotions, while those that get in our way tend to produce anger, sadness, fear, or frustration. Some theories even declare this kind of situational appraisal to be a necessary and sufficient condition for the generation of emotions.²

On closer inspection this picture appears to be too simple. Often we find ourselves in the grip of emotions in which positive and negative qualities are inextricably bound together. The farewell scene in Casablanca is one of the saddest ever to appear on screen. But Rick’s unforgettable gesture, in which he lifts up Ilsa’s head and looks into her eyes, not only expresses deep grief and regret, but also tenderness, pride, determination, and patriotic pathos. Sometimes we wallow in sorrow; some pleasures are embittered with pains. Fears can be pleasant, and joys may be spoiled with regret. Apparently we often confront paradoxical emotions, emotions with neither an exclusively positive nor exclusively negative valence, but which oscillate between these poles. How can this be explained?

At first glance an attractive explanation is that in such cases the valences of different aspects of the emotion-eliciting event diverge. If you indulge your sweet tooth but regret that you thereby violate your diet, you have mixed feelings. While you appreciate the hedonic qualities of your act, you dislike the fact that it doesn’t comply with some (self-imposed) norm. Spinoza advocates views in this direction about the “unsteady soul” (animus fluctuatio) that is plagued by ambivalent emotions. When we simultaneously hate and love something, he argues, the object of our emotion afflicts us with pain or sorrow (tristitia), but at the same time resembles something that usually evokes pleasure or joy (laetitia) in us (Spinoza, *Ethics*, III, 17).³

---

² Cf. for example Lazarus (1991). This also seems to fit our pre-theoretical conceptual schemes. Linguistic studies reported, in all languages that were investigated, positive and negative categories of emotion concepts. See Shaver et al. (1987); Shaver, Wu & Schwartz (1992).

³ “Si rem, quae nos Tristitiae affectu afficere solet, aliquid habere imaginamur simile alteri, quae nos aequae magno Latitiae affectu solet afficere, candum odio habeimus, & simul amabimus.” A similar treatment has more recently been advocated by Patricia Greenspan (1980). Her key example is “friendly rivalry”: if a good friend wins over you in a situation of rivalry, you may feel pleased and at the same time be disappointed. For further distinctions regarding ambivalent emotions see also the helpful discussion in Lenzen (2004).
This account may be useful for some cases, but it cannot deal with others. Spinoza would be hard pressed to explain why people occasionally enjoy indulging in self-pity or are filled with “righteous indignation.” Here it would appear inappropriate to appeal to incongruent appraisals of different aspects of the events that elicit the ambivalent emotion. We feel scandalized by a person’s behavior if it appears morally wrong, or in some other respect normatively inappropriate, to us. Acts of this kind normally lack any aspect that we also appreciate. Yet people’s indignation often does seem to be interspersed with positive feelings, perhaps with tinges of pride or satisfaction with feeling scandalized – as one may see this as an authentic and normatively adequate manifestation of one’s moral values, sense of justice, moral superiority over the perpetrator, and so on.

In such examples, and many others, a phenomenon begins to surface that has not received much attention in the philosophy of the emotions. Emotions, we wish to argue, can be intentionally directed at emotions. We shall call such higher-order emotions meta-emotions. Exploring this phenomenon will not only help us to explain many allegedly ambivalent or paradoxical emotions, but also to analyze complex emotional states where the valence of some higher-order emotion matches that of a lower-level one. Not only do we enjoy certain fears or feel shame about indulging in malicious joy (discrepant valence of first-level and second-level emotion); we are also happy about falling in love or ashamed of being jealous or envious (same valence of first level and second level emotion).

In what follows we argue that meta-emotions of both kinds do indeed exist. We begin with some basic distinctions regarding ordinary first-level emotions and ask how they affect the construal of meta-emotions (section 2). We look at the intentional structure of meta-emotions (section 3), outline the core idea of appraisal theories, and argue that – contrary to first impressions – accounts from this camp support our point that emotions can be intentionally directed at emotions (section 4). Next we consider empirical findings that lend further credence to our model and can be interpreted as offering genetic hypotheses regarding the conditions under which meta-emotions are elicited (section 5). Finally (section 6), we discuss further applications of the concept. In addition to the proposed treatment of certain allegedly paradoxical emotions, our account may also elucidate phenomena such as emotional communication, emotional self-control, privileged access failure for repressed emotions, and survivor guilt.
2. Intentional and non-intentional emotions

Many emotions have intentional objects, others don’t. Let’s call phenomena of the first type (such as John’s fear of a snake, or Martha’s pleasure about her birthday present) “intentional emotions.” A mental state or event is intentional in this sense if it involves, in Brentano’s words, “object-directedness” or a “relation to a content.” As Brentano illustrates, just as in imagination something is imagined and in judgment something is accepted or rejected, “in love something is loved, in hatred something is hated, in desire something is desired, and so forth” (Brentano 1955, book II, chapter 1, section 5). Pace Brentano, however, not all “psychical phenomena” have intentional objects. Neither do all kinds of affective states. Melancholy typically occurs as a kind of emotional background hum without specific intentional contents. Emotional states or processes of this stripe are called “moods.” Some emotion words exclusively denote intentional emotions; others – such as “happiness,” “angst,” “depression,” etc. – can, depending on the context, either refer to intentional emotions or to moods. This distinction does not preclude a mood subsequently picking up intentional objects and, vice versa, what started as an intentional emotion may lose its object-directedness and degenerate into a mood.4

Intentional emotions involve relations between a subject $a$, some emotional state or event $E$, some time $t$ during which $E$ occurs, and some intentional object $o$. The logical form of this relation is $E_t(a,o)$.5 The intentional object may be an individual object or event, or a group of objects or events. You can be frightened by your neighbor’s dog, or by twenty wild boars that suddenly cross your hiking trail. In the latter case the logical form is $E_t(a,\{o_1,o_2, ..., o_n\})$. The object of an emotion may also be a proposition (as when you fear that the boars decide to stampede in your direction). In this case the emotion can be represented as $E_t(a,p)$. Can all other kinds of intentional emotions be reduced to propositional emotions? We don’t think so, but will put this question on the back burner in this essay. Note, moreover, that attributions of emotions do not generally allow

---

4 This has been stressed by Goldie (2000), 17f.; for a detailed discussion see chapter 6. Goldie also argues that the distinction between emotions and moods is rather a matter of degree, and that one should put the difference by saying that emotions “have more specific objects than moods” (17, our emphasis).

5 A similar analysis is proposed by Lenzen (2004), but not with time indices. As we shall see shortly, it is important for our purposes to relate ascriptions of emotional and meta-emotional states and processes to times.
for existential generalization. If Jim fears that Satan possesses his soul, it
doesn’t follow that there is such a thing as Satan (or Jim’s soul). In general,
the statement that \( a \) has, at \( t \), an emotion \( E \) to the effect that \( b \) is \( F \), does
not entail that there actually is something at which \( a \)’s emotion is directed
and that has the property \( F \). Due to this fact, second-order emotions don’t
generally allow for existential generalizations regarding the objects of the
corresponding first-order emotions.

What has been said so far poses two questions: (i) What kinds of
emotion may occur in higher-order forms? (ii) What kinds of emotion can
constitute intentional objects of meta-emotions?

We see no reason to deny that first-level emotions of whatever
type may cause all kinds of secondary affective states, including moods,
agitations, and bodily feelings. But we shall restrict our use of the term
“meta-emotion” to intentional emotions. Since moods lack specific
intentional objects, they cannot \textit{a fortiori} be directed towards objects
constituted by emotions. An emotional state such as “meta-melancholy”
would be a conceptual impossibility in our account. Intentional forms of
joy, fear, guilt, shame, etc., by contrast, can occur as meta-emotions. Typical examples are \textit{angstlust}
or feeling guilty over having engaged in
malicious joy. As for the primary level, moods, (non-intentional) agitations
and so forth can be the \textit{objects} of meta-emotions. For example, you may
hate the feeling of (unspecific, non-intentional) boredom, or be afraid of
falling into a state of depression.

Obviously, emotions can also be directed at other people’s
emotions. You may be glad about your grandma’s happiness that she has
won the lottery. Yet we restrict the notion of meta-emotion to
intrapersonal cases: if \( S \) has a meta-emotion, its intentional object is an
intentional or non-intentional emotion had by \( S \). Note that this restriction
has the consequence that not all emotions are suitable candidates for the
role of meta-emotions. For example, emotions such as envy or jealousy,
which are typically directed towards others and may well become intentional \textit{objects} of one’s feeling guilty or ashamed, will not normally
occur as intrapersonal meta-emotions.

What about “negative” cases? Among the intentional objects of \( S \)’s
emotion can also be the proposition that \( S \) fails to have a certain emotion.
You may feel happy about, proud of, or guilty over having emotionally
tuned out in some situation or with regard to some emotional disposition
you used to have. Since in such cases there is no emotional state or disposition at the primary level, however, we shall again deny such states the accolade of “meta-emotion.” Negative cases may also occur at the meta-level. You may not be happy, or not be grumpy about experiencing (or not experiencing) certain joys or sorrows. Since in such cases as well emotions are absent at one level, these states don’t merit the title “meta-emotions” either.

Mentioning emotional dispositions brings us to one last preliminary distinction. Emotion predicates have dispositional and “occurrent-mental-state” readings. They may refer to (i) emotions that are currently experienced; (ii) complex mental states that stretch over a longer period of time and involve a number of linguistic, behavioral, and physical liabilities; or (iii) personality traits. The latter, too, involve long-term tendencies to think, act, and feel in certain ways. Let us refer only to (ii) as the “long-term disposition” sense of emotion terms (cf. Alston 1967). When Jules says that Jim is enthusiastic (or happy or jealous) he may mean that Jim is currently experiencing these emotions; that at present Jim has strong tendencies to experience them; or that he is generally an enthusiastic (or happy or jealous) kind of person. These distinctions also apply to meta-emotions and their emotional objects. We may currently experience pleasure, annoyance, or fright regarding an emotion that we currently experience. Or we may enjoy or dislike detecting in ourselves some long-term emotional dispositions or character traits. Likewise at the meta-level. For example, one may either currently experience repulsion with respect to “weak” emotions, such as fear or sorrow (as they are currently experienced or thought to lurk around in some dispositional form). Or, we may have long-term dispositions at the meta-level or even harbor personality traits that make us dislike having, or being disposed to have, certain emotions at the primary level. (For example, cynics are often said to have a dispositional aversion to “weak” emotions.)

6 “‘It was really interesting to see what the inside of a stomach looked like, or what human brains looked like,’ police, rescue workers, medical staff, soldiers, and others sometimes say. Later on, they may feel guilty for looking at the body without emotion” (Matsakis 1999, 61).
3. The logical structure of (ascriptions of) meta-emotions

What has been said so far suggests that the general logical form of (an ascription of) a meta-emotion is $E_t(a, A_t'(a))$, where $A$ is an affective state. $E$ may, but need not, be of the same type as $A$, and $t$ may, but need not, be identical with $t'$. Depending on the emotional object $(A_t'(a))$ of the meta-emotion, the latter may also be analyzed in more fine grained ways. The structure of (an ascription of) a meta-emotion the object of which is an intentional emotion is $E_t(a, E_t'(a, o))$. Such meta-emotions can be further broken down into two groups, depending on whether or not the lower-level emotion is propositional. Jim may be annoyed about his fear of the neighbor’s dog, or about his fear that the dog is going to bite him. If the emotional object is propositional the structure of the meta-emotion can be represented as $E_t(a, E_t'(a, p))$.

Does $E_t(a, A_t'(a))$ entail $A_t(a)$, and $E_t(a, E_t'(a, o))$ entail $E_t'(a, o)$? No. The emotional objects of a meta-emotion need not exist. This follows, for example, from the fact that even occurrent emotions have dispositional components, and that we lack privileged access to our mental and behavioral dispositions. A teenager’s pride and pathos of what she believes is her eternal love does not entail that her love is in fact eternal. It doesn’t even entail that it is love at all. Moreover, there is the phenomenon of emotional repression. Repressors tend to misinterpret their emotions or, in certain circumstances, even fail to notice them at all. Meta-emotions are thus not generally veridical.

Nor do we wish to claim that, vice versa, every emotion, in any circumstance, elicits some meta-emotion. Just as $E_t(a, A_t'(a))$ does not entail $A_t(a)$ and $E_t(a, E_t'(a, o))$ does not entail $E_t'(a, o)$, $A_t(a)$ does not entail $E_t'(a, A_t(a))$, and $E_t'(a, o)$ does not entail $E_t(a, E_t'(a, o))$. Our model doesn’t exclude higher-order emotions beyond the second level, but it doesn’t threaten to carry us into infinite hierarchies of emotions. At how many levels, it may be asked, can emotions occur in human (and potentially also non-human) subjects? Can we feel guilty for enjoying the thrills of a bullfight? Or be ashamed of enjoying a good cry during a cheap tear-jerker? It would seem so. Exactly at how many levels meta-emotions may occur is a question that cannot be answered with purely philosophical

---

7 The “negative” cases (which we don’t count as meta-emotions proper) have the form: $E_t(a, \neg A_t(a))$, $\neg E_t(a, A_t(a))$, and $\neg E_t(a, \neg A_t(a))$. 

means, however, and we shall set it to one side here. What has been said so far does, nonetheless, allow for the following general characterization of meta-emotions:

Meta-emotions are occurrent or dispositional intentional emotions had by some subject $S$ at some time (or some period of time) $t$, taking as their objects at least one of $S$’s own dispositional or non-dispositional affective states or processes that occur, or are believed by $S$ to occur, at (or during) $t$ or some (period of) time prior or later than $t$.

It should be evident from this characterization that meta-emotions are, in our view, genuine emotions. They are no sui generis species of mental states. What distinguishes them from first-level emotions is simply the fact that their intentional objects are emotions. But just as higher-order beliefs are, apart from their higher-order structure, “normal” beliefs and higher order volitions, apart from their higher-order structure, “normal” volitions, higher-order emotions are “normal” emotions.

So far we have been looking into the general structure of meta-emotions and their ascriptions. But exactly what kinds of mental entity are meta-emotions, and how are they elicited?

4. Meta-emotions in light of appraisal theories

Until the second half of the twentieth century, both rationalist and empiricist traditions in the philosophy of the emotions were dominated by so-called “feeling theories.” According to feeling theories, the essence of an emotion is its phenomenal quality. Such views face a number of well known problems, however, which have led most authors over the last decades to abandon them. Some authors have reacted by suggesting radical forms of cognitivism in the analysis of emotions. Emotions, they

---


9 Exceptions are Damasio (1994); the theory of “Embodied Appraisals” suggested by Jesse Prinz and his careful defense of the James-Lange theory against some standard objections (Prinz 2003, 2004); and Jenefer Robinson’s thesis that each emotion is the product of “affective appraisals” linked with bodily experiences (Robinson 2004). For helpful discussions on the difficulties that traditional feeling theories face, see the classic investigations of Kenny (1963) and Alston (1967); and Lyons (1980), chapter 1; Deigh (1994); Betzler (2003); and Voss (2004).
claim, can be reduced to evaluative judgments.\textsuperscript{10} If such theories were correct, meta-emotions would simply be judgments (of a certain kind) about judgments (of a certain kind); they would consist in higher order epistemic acts and the resulting attitudes. But reductive forms of cognitivism throw out the baby with the bathwater. For one thing, it does seem essential to the concept of emotion that emotions have phenomenal qualities, even though feeling theories may overemphasize this aspect. Moreover, the judgments and beliefs involved in an emotion are neither necessary nor sufficient for having it. My fear of spiders does not entail that I judge them (whether consciously or not) to be dangerous or in some other way threatening to my well-being. Furthermore, emotions typically involve other propositional attitudes as well, for example volitions. Feelings of shame or sorrow, for instance, also incorporate wishes and desires to the effect that similar situations will not occur in the future. In light of these and other considerations\textsuperscript{11} we favor a view of emotions as mental phenomena with at least five components: (i) phenomenal qualities, (ii) volitions and desires, (iii) evaluative judgments and beliefs, (iv) physiological reactions, and (v) behavioral dispositions, including dispositions to behave in certain linguistic, facial, and bodily ways.\textsuperscript{12} Some of these components (maybe all of them) may also be described in terms of their functional role. It seems plausible to analyze them as systematic relations between inputs to the emotional system, system changes that are caused by these inputs, and corresponding emotional outputs.\textsuperscript{13} It would be beyond the scope of this paper to spell out the details of this model and apply each aspect to the notion of meta-emotion. We shall thus be selective and in the remainder of this essay focus on the cognitive components. What role, if not that envisaged by reductive forms of cognitivism, do judgments and beliefs have for emotions and meta-emotions?

Detailed proposals regarding this role can be found in psychological appraisal theories of the emotions, as developed and

---

\textsuperscript{10} See for example Solomon (1993), 126; Solomon (1988); and Nussbaum (1990) or (2004), 196, where she summarizes her position by claiming that “emotions can be defined in terms of judgment alone.”

\textsuperscript{11} David Rosenthal (1998), for example, argues that emotions differ from propositional acts and attitudes in that the former can be verbally expressed without being conscious.

\textsuperscript{12} A related approach is an integrative definition of emotion proposed by Kleinginna & Kleinginna (1981), which, in psychological contexts, has gained wide currency. For multi-component approaches within philosophical theories, see for example Goldie (2000), or Ben-Ze’ev (2004).

\textsuperscript{13} This has been suggested by Achim Stephan (2004).
empirically tested by Richard Lazarus, Klaus Scherer, Nico Frijda, and others. Such theories are best interpreted as investigating conditions under which emotions are generated. The most important appraisal criteria proposed by appraisal theories include (a) the perceived “novelty” of, or change in, a situation (its “event-likeness”); (b) its (hedonic) “valence”; and (c) its perceived relevance to the subject’s personal ends and goals. In addition, criteria are proposed such as (d) normative adequacy, (e) “agency,” (f) controllability, and (g) what is often called “certainty” regarding the question as to whether or to what degree an emotionally relevant event or situation is or will be present.

The core idea of appraisal theories is that emotions are formed on the basis of characteristic patterns of appraisal. Anger, for example, comes up if situational changes are perceived, and experienced as unpleasant, counter-productive with respect to one’s goals, and normatively inappropriate. Moreover, the subject might feel that s/he or someone else is responsible for the situation and that the situation could be, or could have been, controlled. We may regret that the sun is not shining; such facts too may elicit emotions. But we cannot literally be angry about such kinds of events (unless some deity is perceived to be responsible for them), for they cannot be controlled and don’t have normative dimensions. We may however be angry about the fact that rain is dripping through our recently renovated roof. This event may be perceived as a situational change, as unpleasant, not being in line with our goals and desires, and normatively inappropriate (since we believe the roofer was professionally and legally obligated to do a proper job). Moreover, we may feel that we are in a good position to control and change the situation, e.g. by having the roofer redo his job, or by employing someone else to fix the problem.

Current psychological appraisal theories indirectly support the thesis that emotions can be objects of emotionally relevant appraisals. Consider for example the meta-emotional state of being angry at oneself for one’s fearfulness. In the spirit of psychological appraisal theories this meta-emotion could be reconstructed as follows. The emotion has become salient for the person (situational change/“novelty”); he is convinced that it is an instance of fear, and that it is his fear (certainty, “agency”); he experiences it as unpleasant (hedonic valence), and believes it to obstruct his goals – for example to pass a test of courage, or to make a public appearance that he believes is important for his career (goal relevance). Moreover, he may regard his fear to be normatively inadequate (normative adequacy) and believe that he could overcome it (controllability). An
analogous picture suggests itself for meta-emotions with discrepant first- and second-level valence. If you feel ashamed about having indulged in malicious joy, this may plausibly be explained in terms of the following situational appraisals. You realize the emotion as your joy and acknowledge that you are responsible for it (change/novelty, certainty, “agency”); you acknowledge that the emotion has positive hedonic qualities for you (valence), but regard this to be normatively inadequate (normative adequacy). Furthermore, you believe that your malicious joy does not promote some of your goals, be it the “character goal” never to engage in this kind of emotion or the goal to present yourself as a sensible, empathetic person in the presence of a potential friend or lover (goal relevance). Finally you may also think that you can, at least to a considerable degree, control this emotion (controllability).

These reflections still leave various questions unresolved. The factors just outlined play an important role in current discussions, yet there is scope for debate about how exactly a potential set of necessary and sufficient conditions should be constructed. Does each situation that triggers an emotion of whatever kind involve every single criterion mentioned above? Are they jointly sufficient? Can other forms of appraisal also become a breeding ground for meta-emotions? We must leave these questions for some rainy days, and for future empirical research. The following general conclusion appears justified, however. If situational appraisals do indeed figure in the formation of emotions (which hardly seems deniable), and if this happens only roughly along the lines appraisal theories suggest, many emotions which are prima facie puzzling lose their paradoxical taint. The reason, we suggest, is that both positive and negative emotions can constitute intentional objects of higher-order emotions whose valence diverges from that of the lower-level emotion. Let us see whether we can further fortify this hypothesis with empirical evidence.

5. Empirical findings that suggest the existence of meta-emotions

To date, meta-emotions have not systematically been studied from an appraisal theoretical perspective. There are a number of interesting

14 Seminal works on higher-order affective states, however, are the studies of Peter Salovey, John Mayer and co-workers on the meta-experience of mood, the studies of Mary Beth Oliver on sad film preference, and the studies of John Gottman and co-workers on the influence of parental meta-emotions on child development (cf. Mayer &
investigations, especially in the field of media psychology, which nonetheless suggest that appraisal criteria that have so far only been explored for other (mainly environmental) events and situations can also be applied to emotions.

**Change/novelty.** As already noted, most appraisal theories assume that situations must be perceived as constituting a relevant change in the subject’s environment (i.e. as “event-like”) in order to elicit an emotion. This may with justice be regarded to be the most fundamental step in emotion-related appraisals. The perception of notable changes is supposed to give rise to a state of readiness for emotion. Depending on the outcome of further appraisal processes, this “emotional readiness potential” may or may not develop into a full blown emotion (Ellsworth 1994; cf. also Scherer 1984). Can emotions themselves be perceived as notable situational changes?

Research on “sensation seeking” suggests that strong emotions can satisfy a desire for salient and intense experience. Sensation seeking has primarily been studied in the context of horror films and other portrayals of violence. The strength of the sensation seeking motive has been shown to influence subject’s emotional preferences. Sensation seekers preferred emotionally intense media stimuli, regardless of whether these elicited positive or negative emotions. Participants with a weak sensation seeking motive, by contrast, preferred neutral and positive stimuli (Zuckerman 1979; Zaleski 1984). This suggests that sensation seekers seek experiences of feeling something at all. What some people greet as welcome variations in the monotony of their everyday emotional life, others may perceive as a threat to their emotional equilibrium. In both cases however emotions seem to stand out for their subjects as significant changes in their everyday stream of consciousness.

**Valence/pleasantness-unpleasantness.** Emotions can be pleasant or unpleasant. Feeling exulted is nicer than being grumpy. The critical question for our argument is whether hedonic valence can be regarded as an independent feature of meta-emotional appraisal. Research on “mood management” suggests that it can. According to Dolf Zillmann (2000), for example, the experience of an emotion as being pleasant or unpleasant is a function of the subject’s previous state of arousal. People typically prefer

---

Gaschke 1988; Mayer & Stevens 1994; Salovey et al. 2002; Oliver 1993; Oliver, Weaver, & Sargent 2000; Gottman, Katz, & Hooven 1997).
an intermediate level of arousal and consequently prefer emotions that help them to attain, or remain in, such an intermediate state of arousal. People who are bored and under-aroused enjoy emotionally arousing media stimuli, whereas over-aroused or stressed people prefer soothing media offers. These results suggest that emotions such as fear or suspense can indeed be experienced as being either pleasant or unpleasant, depending on the subject’s previous frame of mind.

**Goal relevance.** Many authors consider the relevance of events and situations for people’s goals, needs, and desires to be the central criterion of emotional appraisal. Can emotions themselves become objects of teleological appraisal? There is empirical evidence that supports this hypothesis. Apparently anger is often welcome in conflict situations because it helps to command respect. In their research on news preferences, Silvia Knobloch-Westerwick and Scott Alter (2006) found indirect evidence for this. Participants of their study were angered by a confederate. One group of participants was then offered an opportunity to pay back the provocateur. Male participants in the “pay back condition” preferred bad news that was likely to sustain their bad moods. As predicted, this effect did not occur in female participants. In the condition without an opportunity to take revenge both sexes preferred good news. This interaction of condition and gender was explained in terms of gender specific norms and social expectations regarding anger and anger expression. Women are not expected to be aggressive, whereas it is often considered “unmanly” not to fight back when provoked. It may be more in accord with the social goals of women, as well as those of men in no-revenge situations, to dissipate anger. But in the male/pay back condition, anger was goal-conducive in the experiments.

According to leading appraisal theories, application of the three appraisal criteria considered thus far provides a sufficient condition for the formation of emotions. The point of the above considerations was to demonstrate that there is empirical evidence that people do indeed apply these criteria to emotions, at least under certain experimental conditions. Whereas these first three criteria are regarded as constituents of all emotions, further appraisal criteria, like “agency,” “control,” “certainty,” and “norm compatibility,” are often thought to be responsible for the differentiation of specific emotions. Is there also evidence to the effect

15 Cf. Zuckerman (1979) and Zuckerman (1994), 27, where he characterizes sensation seeking as “the seeking of varied, novel, complex and intense sensations and experiences, and the willingness to take ... risks for the sake of such experience.”
that meta-emotions have appraisal content that goes beyond a simple positive-negative evaluation of first-order emotions?

*Normative appraisal.* We can, and actually do, appraise emotions vis-à-vis their social and moral propriety. There is a long research tradition, pioneered by Paul Ekman, which deals with culturally defined display rules for emotion (Ekman & Friesen 1969; Ekman 1972; Buck 1984). Furthermore, Arlie Hochschild has shown that not only the external expression but also the inner experience of emotions, can become an object of normative appraisal (Hochschild 1983; cf. also Averill 1991; and Planalp 1999). Hochschild coined the term “feeling rules” to refer to the normative aspect of subjective emotional experience. Feeling rules have been studied most frequently in a professional context. They can be explicitly laid down in “company philosophies,” as in the case of flight attendants or Disneyland workers, who are expected never to get angry with their customers, no matter what happens (Hochschild 1983). In other cases, feeling rules are “unwritten laws,” as with construction workers, among whom fear of accidents – not only talk about accidents, or behaving as if one is afraid – is a taboo (Haas 1978). In both cases, emotions are subjected to normative appraisal, concerning not only their outward expression, but also their private experience.

What about the criteria “agency,” “controllability,” and “certainty”? The agency criterion concerns the question who or what caused the emotion-eliciting event. It seems plausible to assume that emotions are also evaluated in this respect. A person who watches a horror film is normally cognizant that doing so will cause emotions such as fear, horror, or disgust in her, and that she is responsible for creating, or subjecting herself to, a situation that is likely to evoke such emotional experiences. We are not aware, however, of any empirical studies concerning people’s sense of causal agency in emotion elicitation.

There is ample evidence, however, that people appraise emotions with respect to their controllability. A great number of conscious and unconscious strategies for coping with emotion have been studied in this regard, e.g. by Richard Lazarus.16 Finally, emotions would seem to be appraised also with respect to the subject’s certainty about the situation. Sometimes we are unclear, initially at least, about our feelings. Are we tired

---

16 See, for example, Lazarus (1991). Lazarus himself, however, does not consider the person’s sense of control over his or her emotions as a second-order criterion of appraisal, but subsumes it under the first-order appraisal of control over the emotion-eliciting situation.
or depressed, angry or jealous, driven by admiration or infatuation? John Mayer and coworkers developed a “meta-mood-scale” that measures, among other things, the extent to which people think they are clear about their feelings. Participants could report on this scale, in reliable patterns, how certain they were about their feelings (Mayer & Gaschke 1988).

We have argued that each of the appraisal criteria which are widely thought to produce emotions when applied to non-emotional events and situations, can be applied to emotions themselves. We cannot provide conclusive arguments for this thesis. Yet there is considerable empirical evidence that the criteria in question are applied to emotions, at least under certain experimental conditions.17 Irrespective of the specific set of criteria discussed above, it seems justified in light of what has been said to supplement our characterization of meta-emotions with the following thesis about the generation of meta-emotions:

Meta-emotions are elicited when a person appraises his or her own emotions in light of emotionally relevant appraisal criteria.

6. Further applications: emotional self-awareness, art and entertainment, survivor guilt

Our account of meta-emotions sheds light on a number of perennial questions from the philosophy and psychology of emotions and the philosophy of mind. We already have seen how it bears on the problem of allegedly paradoxical emotions. In the final pages of this essay we want to put some more flesh on our proposal by looking at three further examples: emotional self-awareness and self-regulation, emotions in art and entertainment, and survivor guilt.

Self-awareness and self-control of emotions. In seminal work of the psychologist John Mayer and co-workers (cf. Mayer & Gaschke 1988), the concept of meta-emotion is closely related to the notion of “emotional intelligence.”18 Emotional intelligence is construed by these authors as a person’s capacity to experience and express emotions in self-conscious and self-controlled ways. It would be worthwhile to elaborate this idea further and to analyze the exact function of meta-emotions in self-awareness and

17 A more comprehensive overview of research literature supporting the assumption of second order appraisals in the context of media psychology can be found in Bartsch et al. (2006).
self-control of emotions. For instance, it seems intuitively plausible that meta-emotions influence people’s willingness and motivation to pay conscious attention to their first-order emotional states. Emotions that are judged to be enjoyable, goal conducive, morally adequate, or otherwise desirable should be more welcome as objects of self-awareness and self-reflection than emotions associated with pain, failure, moral scruples, and so on. In our view, research on emotional repression provides empirical evidence for this role of meta-emotions in emotional self-awareness. As we have discussed in more detail elsewhere (Jäger & Bartsch 2002; 2006), people’s tendency to repress emotions – i.e. to avoid conscious awareness of emotional arousal – depends on the perceived threat that the emotions pose to their self-esteem (cf. Mendolia 1999).

Meta-emotions have also been related to other forms of emotional self-regulation. John Gottman and coworkers (Gottmann, Katz & Hooven 1997) concluded from their research on “emotional communication” in families that positive meta-emotions help to regulate negative first-order emotions. People with positive meta-emotions about anger and sadness feel rarely out of control with these emotions, compared to people with corresponding negative meta-emotions. Children who grow up in families with a generally positive attitude towards emotions are able to recover more quickly from physiological arousal, are less vulnerable to negative daily moods, and engage in cooperative peer interaction more often than children who grow up in emotion avoiding or emotion dismissing families. If such functions of meta-emotions in self-awareness and self-control of emotions could be further substantiated this would have interesting implications for theories of self-consciousness and personhood.19

Art and Entertainment. Another field where the concept of meta-emotion can be fruitfully applied is psychology of entertainment and art appreciation. Both art and media entertainment often revolve around emotions that are not desired by most people in everyday life. What Mary Beth Oliver (1993) called the “paradox of the enjoyment of sad films” pertains not only to the enjoyment of tear-jerkers, but also to people’s taste for horror films, thrillers, and a variety of other genres renowned for

18 Mayer and Gaschke also talk about “meta-moods”; this does not refer to intentional meta-moods however, but to what they call the “meta-experience” of moods.
19 Another interesting question in this context is the relation between meta-emotions and empathy, and the significance of this relation for altruistic behavior. Empathy has emotional dimensions, and emotions about other people’s emotions may influence one’s intra-personal meta-emotions, and vice versa. For a helpful discussion of the emotional aspects of empathy and their significance for altruistic motivation see Sherman (1998).
eliciting strong negative emotions in their recipients. Oliver proposed to solve this paradox by invoking meta-emotions. She was able to show that feeling sad is central to what drama-fans appreciate about sad films. Oliver observed that in participants with a preference for dramas the reported degree of sadness while watching sad film segments was proportional to the reported degree of enjoyment. Comparable results were found in a study on horror films that one of us conducted. Participants liked horror films more if they reported that they liked the emotions they experienced during the film. Why then do people enjoy and appreciate emotions like sadness, fear, and disgust in the context of media use, but rarely so in everyday life? It can be speculated that media entertainment is designed in such a way as to invite positive meta-emotions by offering a wide range of emotion-related gratifications. Entertainment media foster a state of mild and pleasant arousal, or provide viewers with “kicks” of extraordinarily strong emotion. They make emotions intelligible and controllable for the viewer, and offer moral justifications for the experience and expression of emotions. Conventions of typical entertainment genres can to a large extent be understood as implicit “viewing contracts” that make the gratification potential of genre-typical emotions predictable for the viewer.

**Survivor guilt.** Another interesting application of the concept of meta-emotion is people’s apparently irrational feeling of severe guilt connected with events for which they bear no, or only little, responsibility. A striking example is (the experience of) survivor guilt. Survivor guilt comes in different varieties. According to leading authors in the field, so-called “existential survivor guilt” is an amalgamation of guilt and related feelings about staying alive in situations where others died, or about being less harmed than other people with whom the subject identifies. As the American psychologist and psychotherapist Aphrodite Matsakis characterizes it, so-called *existential survivor guilt* is “the guilt you feel because you are alive, healthier, or otherwise better off than someone else” (Matsakis 1999, 40, 69). *Content survivor guilt*, by contrast, is guilt resulting from something you did to stay safe or alive, including guilt over adopting coping mechanisms such as denial, repression, rationalization, lying, and so forth, or “having certain thoughts and feelings during trauma or stressful ordeal” (Matsakis 1999, 58; cf. also Williams 1987). The boundaries

---

20 Results from an unpublished study conducted by Anne Bartsch and Dennis Storch (Bartsch & Storch, in preparation).
21 This has been suggested by Bartsch et al. (2006), and Bartsch and Viehoff (2003).
between the two forms are fuzzy, since the fact that a person survived a tragic event may have resulted from something s/he did. But this need not detain us here. To keep matters within reasonable bounds in the present context, we shall focus on paradigm cases of existential survivor guilt. Let’s narrow down the territory further by considering a particular kind of survivor guilt, namely what may be called existential survivor guilt proper, i.e. guilt over having survived in a situation where others died. This kind of survivor guilt may again be broken down into two forms.

(i) In some cases the person who feels guilty played a role in the tragic incident. A good example of this is the case of Wan Dang, one of the protest leaders of the 1989 Tiananmen Square demonstrations for democracy in China. When hundreds of demonstrating students and other Chinese civilians were shot by the Chinese army, Wan Dang felt, as he later confessed, guilty for this incident – despite his being fully aware that “the main responsibility was not ours [but] with the government who actually did the killing” (The Sun, 1998, A3, quoted from Matsakis 1999, 53). In such forms of survivor guilt, which are connected with “role related responsibility guilt,” the extent of the experienced guilt is typically not warranted by the subject’s actual responsibility. Still, such feelings may not be entirely inappropriate. Small as it may have been, the subjects did have their share in the causal chains that led to the tragic events, and self-assessments of the correct degree of responsibility are characteristically difficult in such cases (especially for those who were involved). We shall therefore turn to another type of existential survivor guilt, which appears even more puzzling.

(ii) Often survivor guilt is experienced in connection with situations that have nothing whatever to do with one’s own actions, and in which one knows that this is so. This may be called “existential survivor guilt proper without perceived responsibility,” or for short: “survivor guilt without responsibility.” It is especially this phenomenon, we suggest, in which meta-emotions play a crucial role. Consider survivors of natural catastrophes, car or ship accidents, plane crashes, and so forth, who know they couldn’t have done anything about the incident and couldn’t have saved a single life. Often such people still experience survivor guilt. If the intentional objects of these feelings were the events themselves, or the fact that the survivors had been spared, their emotions would appear entirely irrational. But, as has recently been argued by David Velleman (2004), such

---

22 We shall henceforth often drop this qualification and use the term “survivor guilt” in a sense that doesn’t imply that the subject actually is guilty of anything.
a diagnosis would leave something out. Feelings of guilt without perceived wrongdoing, Velleman insists, can be rational and warranted. There is much to be said for this diagnosis. Yet Velleman’s explanation, while it points in the right direction, may be found less convincing.

According to Velleman, the fact that you haven’t wronged anyone doesn’t imply that you have no grounds for feeling guilty. Instead, “it may only show that we need to interpret our feelings [of guilt] more carefully, as anxiety about warranting envious resentment” (Velleman 2004, 247f.). There can be grounds for envying those who are more fortunate, Velleman thinks, “or for resenting those whose good fortune is undeserved; and so ... a beneficiary of good fortune may rationally feel anxiety about providing others with ground for resentment” (ibid.). Now, the envious resentment that survivors fear obviously cannot be felt by those who were (maybe literally) in the same boat, but did not survive. Velleman thus argues that third parties can feel “resentment on behalf of the deceased. ... And a survivor can rationally feel anxiety about providing grounds for such vicarious or sympathetic resentment” (ibid.). However, can this form of anxiety really be construed as (an experience of) guilt? Even if it can, would not the subject have to believe that s/he provides good grounds for resentment? And why should this be so? An explanation in terms of meta-emotions, we believe, provides a more plausible and simpler explanation of survivor-guilt-without-responsibility. In Velleman’s account, too, this peculiar kind of guilt is explained as an emotion about emotions. But Velleman construes the phenomenon as an interpersonal form of higher-level emotion. The truth, we contend, is that survivor guilt is an intrapersonal meta-emotion. Survivors experience positive emotions about their fortune (such as gratitude and relief). But at the same time they feel guilty about having these positive emotions. In light of the tragic losses the event caused for others, the survivors feel – consciously or unconsciously – that any positive emotions related to the situation are (normatively) inappropriate. According to this view, survivor guilt is not a special form of anxiety, but genuine guilt. However, it is not (the experience of) guilt over some “wrongdoing” in the ordinary sense, but guilt over being swamped with positive feelings despite the fact that grief and sorrow alone are perceived to be appropriate.

This explanation fits nicely into the picture psychologists and psychotherapists draw of the phenomenon. For example, Matsakis explicitly points out, on the basis of extensive experience with post-traumatic clients, that “survivor guilt does involve gratitude” (1999, 50).
She reports the story of Paul and Karen whose car was struck by a truck when they were driving home one night. Karen was severely injured; Paul was not.

“Paul wondered why he was spared injury while his wife had to endure numerous surgeries and much pain. ‘It should have been me. And if it couldn’t have been me, why couldn’t the injuries at least have been divided equally between us?’ Paul asked others. Paul meant what he said. ... On the other hand he was glad he had been spared Karen’s ordeal. ... Paul felt guilty about being grateful he wasn’t hurt, just as guilty as he felt about not having been injured” (Matsakis 1999, 50).

This is not an instance of “existential survivor guilt proper,” as we dubbed it (for Karen survived); but naturally people also feel glad for having been spared from death, instead of feeling gratitude for not suffering from some (non-lethal) injury.

The psychological dispositions of survivors may also elicit meta-emotions about positive experiences and emotions that are not directly related to the survival. As in the case of Mitchell: “The tornado destroyed almost every home in the neighborhood except mine,” he reported. “Six people died. The people next door lost their baby and everything they owned. The night after that tornado hit our town, my wife and I had the best sex we’d had in years. Is there something wrong with me? Why am I glad I’m okay when others around me are hurting?” (Matsakis 1999, 39). Here Mitchell doesn’t talk about guilt he feels for being glad that he has been spared, but for enjoying pleasures and positive emotions while others he feels close to are thrown into extremely desperate and painful situations. Speaking more generally, survivor guilt may involve the feeling that “you should not go on with life, or at least not enjoy success or happiness” (Matsakis 1999, 43). What started as a negative meta-emotion towards positive emotions directly experienced in connection with one’s survival (like gratitude and relief) may in such cases develop into more general negative meta-emotions towards enjoying positive emotions at all. However, the crucial point in all these cases, and those discussed above, is that we are clearly dealing with emotions whose intentional objects are other emotions.

Let us sum up. We have presented a number of phenomenological and ordinary language considerations that suggest the existence of meta-
emotions. We have looked into the conceptual nature of the latter, and outlined the basic structure of ascriptions of meta-emotions. Next, we have squared our thesis with psychological research on emotion and, drawing on specific empirical findings, formulated a general genetic hypothesis about the conditions under which meta-emotions arise. In this context, our focus was on appraisal theories. How, it may be asked, can meta-emotions be treated within other approaches? An answer to this question must wait. But, in light of what has been said in this essay, the following conclusion appears justified: Whatever format an acceptable theory of emotions adopts, it should be equipped to deal with meta-emotions. This is a constraint for any theory. Apart from the suggested treatment of certain ambivalent emotions, the applications discussed in the present section show that meta-emotions are not merely a philosophical construction. We are dealing with reality here.

REFERENCES


23 Bartsch (forthcoming) examines how meta-emotions might be treated in neurological, script-theoretical, and social constructivist accounts of emotion.

24 For valuable comments on earlier versions of this paper we wish to thank Nina P. Azari, Christian Beyer, Dieter Birnbacher, Carl Plantinga, an anonymous referee for the Grazer Philosophische Studien, and especially Darrell Rowbottom.


James, W., 1884: “What is an Emotion?”, Mind 19, 188-204.


Salovey, P., L.R. Stroud, A. Woolery, & E.S. Epel, 2002: “Perceived Emotional Intelligence, Stress Reactivity, and Symptom Reports: Further Explorations Using the Trait Meta-Mood Scale”, *Psychology and Health* 17, 611-627.


University of Chicago Press, 405-434.


