

What is an Emotion in the Belief-Desire Theory of Emotion?

Rainer Reisenzein¹

University of Greifswald, Germany

To appear in: F. Paglieri, L. Tummolini, R. Falcone & M. Miceli (Eds.), *The goals of cognition: Essays in honor of Cristiano Castelfranchi*. London: College Publications (2012, in press).

Abstract Let us assume that the basic claim of the belief-desire theory of emotion is true: What, then, is an emotion? According to Castelfranchi and Miceli (2009), emotions are mental compounds that emerge from the gestalt integration of beliefs, desires, and hedonic feelings (pleasure or displeasure). By contrast, I propose that emotions are affective feelings caused by beliefs and desires, without the latter being a part of the emotion. My argumentation for the causal feeling theory proceeds in three steps. First, I argue that affective feelings should be regarded as components of emotions because this assumption provides the best available explanation of the phenomenal character and the intensity of emotional experiences. Second, I examine the two main arguments for regarding beliefs and desires as emotion components—that doing so is needed to explain the finer distinctions among emotions and their object-directedness—and argue that they are unconvincing: Emotions can be distinguished by referring to their cognitive and motivational causes, and their appearance of object-directedness could be an illusion. Third, I present three objections against the hypothesis that beliefs and desires are components of emotions: This hypothesis fails, at second sight, to explain the directedness of emotions at specific objects; it has difficulty accounting for the duration of emotional reactions caused by the fulfillment of desires and the disconfirmation of beliefs; and there are reasons to question the existence of the postulated emotional gestalts and the process that presumably generates them. The causal feeling theory avoids these problems. I therefore recommend abandoning the belief-desire compound theory of the nature of emotions in favor of the causal feeling theory. However, a partial reconciliation of the two theories is possible with respect to the concept of “affectively tinged” thoughts.

¹ I thank Maria Miceli and Martin Junge for their comments on a previous version of the manuscript. Address all correspondence to Rainer Reisenzein, Institute of Psychology, University of Greifswald, Franz-Mehring-Strasse 47, 17487 Greifswald, Germany. E-mail: rainer.reisenzein@uni-greifswald.de

1. The Belief-Desire Theory of Emotion

1.1 Basic Assumptions of BDTE

The task of emotion psychology is to develop an accurate, reasonably detailed and comprehensive model of the human emotion system, including its interactions with other subsystems of the mind. I believe that, of the different theoretical approaches to emotion, the cognitive approach holds the greatest promise for attaining this goal. This belief is shared by many of my fellow psychologists, which partly explains why cognitive theories have dominated the psychological discussion of emotions during the past 30 years (e.g., Frijda, 1986; Lazarus, 1991; Oatley & Johnson-Laird, 1987; Ortony, Clore & Collins, 1988; Scherer, 2001; review in Ellsworth & Scherer, 2003). The same is true for philosophy (e.g., Lyons, 1980; Roberts, 2003; Solomon, 1976; review in Goldie, 2007). Likewise, cognitive theories of emotion dominate artificial intelligence research on emotions (e.g., Marsella, Gratch, & Petta, 2010).

However, the cognitive approach to emotions is not homogeneous. Rather, there are different cognitive emotion theories, and some are more plausible than others (see e.g., Reisenzein & Döring, 2009). In this article, I restrict my attention to what I regard as the most plausible version of cognitive emotion theory: the *cognitive-motivational*, or *belief-desire theory of emotion* (BDTE). Originally proposed by philosophers (e.g., Davis, 1981; Green, 1992; Marks, 1982; for an early version, see Meinong, 1894 [summarized in Reisenzein, 2006]), BDTE is attracting increasing interest in psychology and artificial intelligence research (e.g., Reisenzein, 2001a; 2009a; 2009b; Mellers, 2000; Marsella et al., 2010). Writing at the intersection of these areas, Maria Miceli and Cristiano Castelfranchi have been long-standing proponents of BDTE (e.g., Miceli & Castelfranchi, 1997; 2002; 2007; Castelfranchi & Miceli, 2009).

Seen from up close, BDTE itself turns out to be a family of theories rather than a single, uniform theory. What unites the members of the BDTE family is a basic assumption about the “psychological preconditions” (Meinong, 1894) of emotions. The assumption can be formulated as follows: A central subset of the mental states presystematically regarded as emotions—roughly, those that seem to be directed at propositional objects (i.e., actual or possible states of affairs)—presuppose, for their existence, beliefs and desires concerning these objects.² Beliefs and desires, in turn, are regarded in BDTE as basic kinds of representational mental states that cannot be reduced to one another: Beliefs aim at truth and have a cognitive or information-providing function, whereas desires aim at satisfaction and have a motivational function (Green, 1992, p. 18).

To illustrate, according to BDTE, Mary feels *happy* that *Mr. Schroiber was elected chancellor* (only³) if she comes to believe that this state of affairs *p* is the case, and if she desires *p*; whereas Mary feels *unhappy* that Schroiber was elected chancellor if she is

² Strictly speaking, this is true only for emotional reactions to events considered to be real (Meinong [1910] called them “serious emotions”), not for emotional reactions to fictions (Meinong’s “fantasy emotions”). However, BDTE can be extended to deal with fantasy emotions, by replacing *beliefs* with *assumptions* (Meinong, 1910; Reisenzein, Meyer, & Schützwohl, 2003; Reisenzein, 2012).

³ As far as I can see, all BDTE theorists assume that under normal conditions (roughly, in the normally functioning, awake adult), the co-occurrence of the beliefs and desires required for an emotion is also, at least *causally*, sufficient for the occurrence of that emotion. Given normal circumstances, the “if” in the following if-then laws of BDTE can therefore be read as “if and only if” rather than just as “only if.”

averse to p or “diswants” p to happen (which I shall here analyze as: she desires not- p), and comes to believe that p is the case. Hope and fear can be analyzed by allowing for beliefs held with uncertainty: Mary *hopes* for p if she desires p but is uncertain about p (i.e., her subjective probability that p is the case is between 0 and 1), and she *fears* p if she desires not- p and is uncertain about p . Several other emotions can be brought into the scope of BDTE if one takes the experiencer’s pre-existing beliefs into consideration (Reisenzein, 2009a): Mary is *surprised* that p if she up to now believed not- p and now comes to believe p ; she is *disappointed* that not- p if she desires p and up to now believed p , but now comes to believe not- p ; and she is *relieved* that not- p if she is averse to p and up to now believed p , but now comes to believe not- p . Social and moral emotions such as joy and pity for another, or guilt and moral elevation, can be analyzed in BDTE by introducing other-regarding desires—desires that concern the fate and actions of other agents (see Reisenzein, 2010; and Castelfranchi and Miceli, 2009; Miceli & Castelfranchi, 2007).

Although the belief-desire theory of emotion has been around for some time, it has so far been only partially explored, both theoretically and empirically. At least it has been much less explored than the belief-desire theory of action, BDTE’s cousin in the domain of goal-directed behavior. Just as BDTE assumes that beliefs and desires are the mental preconditions of emotion, the belief-desire theory of action assumes that beliefs and desires—in this case, beliefs about means-ends relations, and desires for the ends—are the mental preconditions of action (e.g., Conte & Castelfranchi, 1995). The belief-desire theory of action has, under a variety of different names and in several conceptual disguises, become the object of an extensive research program in several disciplines spanning philosophy, psychology, sociology, and artificial intelligence (for a selective sample of this literature, see e.g., Bratman, 1987; Conte & Castelfranchi, 1995; Feather, 1982; Fox & Poldrack, 2009; Goldthorpe, 1998; Mele, 1992; Reisenzein, 2001b; Wooldridge, 2002). No comparably extensive research program exists to date for BDTE, although several explorations of the theoretical domain have been made (e.g., Castelfranchi & Miceli, 2009; Davis, 1981; Green, 1992; Mellers, 2000; Miceli & Castelfranchi, 1997; Reisenzein, 2009a; 2009b; 2010; Reisenzein & Junge, 2012). As Castelfranchi and Miceli (2009) note, further exploration of BDTE is needed.

1.2. Emotions: Belief-and-desire-caused Feelings or Compounds of Beliefs, Desires, and Feelings?

In this article, I focus on one of the issues that divide the members of the BDTE family: the question of what an emotion is in BDTE. The answer to this question—the question of the nature of emotion—is intimately connected to the interpretation of the basic claim of BDTE, that beliefs and desires are necessary preconditions of emotions (see also, Reisenzein, 1994a; 2000). Three main proposals regarding the more precise relation between beliefs/desires and emotions, and corresponding to these, three main proposals regarding the nature of emotion in BDTE have been made. Illustrated for the case of happiness about p , they are as follows.

1. The *causal view* holds that the belief that p and the desire for p are (necessary) causes of happiness about p , whereas the emotion is a separate mental state, such as a feeling of pleasure or displeasure (e.g., Reisenzein, 2009a).

2. The *part-whole view* claims that the belief that p and the desire for p are not, or at least not only, the causes of the emotion, but parts of the emotion. Correspondingly, the emotion is conceived of as a complex mental state consisting of the co-occurrence of (2a) the belief that p and the desire for p (e.g., Marks, 1982); or (2b) the belief and desire plus other components (e.g., a feeling of pleasure or displeasure).

3. The *fusion view* maintains that happiness about p is a new mental state that emerges through a process of “fusion” (Green, 1992) or mental integration from (3a) the belief that p and the desire for p (Green, 1992) or from (3b) these two and further components (e.g., feelings of pleasure or displeasure; Castelfranchi and Miceli, 2009).⁴ Characteristic of the fusion view is not only the assumption that the output of the fusion process—the emotion—is a phenomenally unitary mental state; but also the assumption that this state has emergent properties, such as a unique experiential quality (Castelfranchi & Miceli, 2009; Green, 1992).

In this article, I compare two specific versions of the causal view and the fusion view of the nature of emotions in BDTE: the *causal feeling theory* proposed by Reisenzein (2009a; 2009b), and the *belief-desire compound*, or *emotional gestalt theory* proposed by Castelfranchi and Miceli (2009).

1.2.1 The Causal Feeling Theory

Reisenzein (2009a) proposed a computational (C) model of the belief-desire theory of emotion called CBDTE, according to which emotions are the products of two hardwired comparator mechanisms that service the belief-desire system, the belief-desire comparator (BDC) and the belief-belief comparator (BBC). These mechanisms constantly compare, at an unconscious level of information processing, newly acquired beliefs about the world or the self with, respectively, pre-existing desires (BDC) and beliefs (BBC). If a match (p ; p) or a mismatch (p ; not- p) between the contents of a new belief and those of an existing belief or desire is detected, the comparator mechanisms generate nonpropositional, sensation-like output signals that communicate the information about the detection of the match or mismatch to other cognitive subsystems. Based on this computational explication of BDTE, I then proposed that emotions should be identified with the output signals generated by the BDC and the BBC (see also, Reisenzein, 2009b). This proposal implies that emotions, although caused by certain belief-desire constellations, are distinct from the latter and do not contain them as components. Specifically, emotions are nonconceptual and nonpropositional mental states that are, when conscious, experienced as feelings of pleasure and displeasure, surprise and expectancy confirmation, combinations of these feelings (e.g., disappointment, relief), and hope and fear. A main reason for proposing this theoretical (theory-

⁴ Philosophers typically classify the causal version of BDTE as a “noncognitive” theory of emotion. The reason is that their criteria for classifying emotion theories are the theories’ assumptions about the nature of emotion. The causal theory, although assuming that emotions presuppose cognitions as their causes, takes the emotion itself to be a noncognitive mental state—at least in the sense that it is not, and does not include, a belief. (Note, however, that this still allows the emotion to be cognitive in the wider sense of being a representational mental state). Theories of types 2b and 3b are sometimes called “hybrid” theories of emotion, because they assume that emotions contain both cognitive and noncognitive components (e.g., Reisenzein, 1994a). By contrast, theories of types 2a and 3a are called “pure” cognitive (or cognitive-motivational) theories of emotion. For further discussion, see e.g., Green (1992), Prinz (2004), Reisenzein (1994a), and Reisenzein and Döring (2009).

based; see Section 2) definition of emotion—the identification of emotions with the output signals of the BDC and BBC—was causal-functional: These signals recommended themselves as the scientific referents of the term “emotion” because, assuming that CBDTE is a correct description of the emotion-generating mechanisms, they constitute the “causal hub in the wheel of emotion” (Reisenzein, 2009a): They are proximately caused by beliefs and desires (the inputs of the emotion mechanisms) and they are in turn (partial) causes of all emotional effects postulated in the theory—emotional experience, shifts of attention, updates of the belief-desire system, emotional actions, expressions, and physiological changes. In addition, I argued that the proposed theoretical definition of emotion provides a natural explanation of the salient properties of emotional experiences, in particular their phenomenal quality and intensity (see also, Reisenzein, 2009b).

1.2.2 The Belief-Desire Compound (Emotional Gestalt) Theory

In contrast to the causal feeling theory, Castelfranchi and Miceli (2009) propose that emotions are mental *gestalts* in the sense of the Gestalt theorists (e.g., Köhler, 1947). As the authors explain, these emotional gestalts

“result from the ‘fusion’ [Gestalt integration] of the BD [belief-desire] compound with affect (feelings of pleasure or displeasure). In this process, the BD compound is hedonically colored by affect, while the latter is specified and qualified by the BD compound. The emotional experience resulting from this fusion is...an emergent property which cannot be traced back to any single component of the emotion” (Castelfranchi & Miceli, 2009, p. 228).

To illustrate this theory, Mary’s happiness that Schroiber was elected chancellor would result from the fusion, or gestalt integration, of Mary’s belief that this state of affairs p is the case, her desire for p , and the pleasant feeling generated by the co-occurrence of these two mental states. According to Castelfranchi and Miceli (2009), the resulting emotional gestalt, although being a causal effect of the belief that p and the desire for p , can also be said to contain these mental states as parts.⁵ Thus, the gestalt theory of emotion combines assumptions of the causal and the part-whole theory. It may be noted that the gestalt theory represents one of two existing explications of the idea that emotions result from a fusion or integration of several components. The other explication of this idea holds that the fusion process is a process of categorization in which the different components of an emotion instance or token are subsumed under an emotion schema (e.g., Mandler, 1984; Barrett, 2006). For discussions about this alternative, see Castelfranchi and Miceli (2009) and Reisenzein (1994a).

My aim in this article is to compare the two described theories concerning the nature of emotion in BDTE, the causal feeling theory and the emotional gestalt theory. This comparison is facilitated by the fact that, apart from their different assumptions about the nature of

⁵ At first sight, a component α of a complex mental state $\gamma = (\alpha \ \& \ \beta)$ cannot also be the cause of γ , for that would imply self-causation, something generally regarded as impossible (see e.g., Mackie, 1974; Reisenzein & Schönplflug, 1992). Castelfranchi and Miceli (2009, p. 229) concede this point but argue as follows: “the BD [belief-desire] compound...(α)...together with affect (β) causes the emotional gestalt (γ)...when α is included into γ , it is no longer there as such; rather, in virtue of its being part of the whole (γ), it changes into α' . This is analogous to a line segment that, when it merges with other lines into the gestalt of a triangle, is no longer perceived as a (mere) line segment but as a side of the triangle. If this is accepted, one can legitimately talk about the BD compound as being *both* a cause (α) and a constituent (α') of emotions.”

emotion, the versions of BDTE endorsed by Castelfranchi and Miceli and myself are close. Most important for the following discussion, both versions of BDTE assume that “emotional” belief-desire configurations (those considered to be necessary for an emotion; e.g., in the case of happiness about p , the desire that p plus the belief that p) cause feelings of pleasure or displeasure, and that these hedonic feelings are an essential component of emotions (in CBDTE: for all emotions but hedonically neutral surprise). Furthermore, we agree that the computational processes that underlie the causal link between belief and desire on the one hand, and hedonic feelings on the other hand, involve a comparison of belief and desire (on this point, see also Miceli & Castelfranchi, 1997). However, whereas I propose to identify emotions with the output signals generated by the belief-desire comparator—as well as the belief-belief comparator, which is not explicitly considered a part of the emotion mechanism by Miceli and Castelfranchi—Castelfranchi and Miceli (2009) identify the emotion with a mental gestalt that integrates belief, desire, and hedonic affect. These differences in the theoretical definition of emotion imply another, causal or processual difference between our respective versions of BDTE: Whereas Castelfranchi and Miceli are committed to the existence of the hypothesized gestalt formation process and its outcomes, the emotional gestalts, CBDTE assumes neither a gestalt formation process nor emotional gestalts.⁶ Castelfranchi and Miceli’s version of BDTE is therefore, in this respect at least, more complex than CBDTE.

2. What is a Definition of Emotion?

Before proceeding, I need to address an antecedent issue. This issue concerns the metatheoretical status of the problem under discussion, the question “What is an emotion in BDTE?”

The issue addressed by the question “What is an emotion?” is traditionally called the problem of the *nature* of emotion. Alternatively, it is called the problem of the *definition* of emotion (e.g., Scherer, 2005). I will use the latter formulation because it is more general. As usually understood, a definition is a specification of the essential or necessary (or, if such don’t exist, at least the typical) features of a class of objects. Since Aristotle, two different bases of definitional necessity have been distinguished: linguistic conventions and the structure of language-independent reality. These two bases of necessity are reflected in the traditional distinction between nominal and real definitions (e.g., Boyd, 2002). As traditionally understood, *nominal definitions* reflect conventions about how a term is to be used, whereas *real definitions* capture the pre-existing essential properties of a class of objects.

In terms of this distinction, the question “What is an emotion?” seems to be *prima facie* asking for a real definition: a description of the nature of emotions, their pre-existing essential features. And this is how the question has been understood by most classical (e.g.,

⁶ Apart from the question of the nature of emotion, there are several other differences between our respective versions of BDTE. For example, Miceli and Castelfranchi (1997) assume that at least for some kinds of emotions, the belief-desire comparison process involves meta-beliefs about the presence of a discrepancy between belief and desire, whereas I reject this idea (Reisenzein, 2009a). I distinguish between emotions and emotional experience, whereas Castelfranchi and Miceli largely equate the two. And whereas I assume that the emotions covered by BDTE require a propositional representation system (a language of thought; Fodor, 1987), Castelfranchi and Miceli (personal communication) find this assumption too restrictive and prefer a more liberal view of the format of the mental representations that may underlie the emotions covered by BDTE. In this article, I will ignore these differences because they are not decisive for the present discussion of the nature of emotions in BDTE.

James, 1890/1950; Meinong, 1894; Wundt, 1896) and many contemporary emotion theorists (e.g., Arnold, 1960; Schachter, 1964), at least in psychology.⁷ However, during the heyday of logical empiricism, the view became popular that the concept of real definition is obscure and unscientific (e.g., Hempel, 1965). For the adherents of this view, the question “What is an emotion?” and parallel questions about specific emotions, such as “What is fear?”, which *prima facie* seem to ask for real definitions, are either nonsensical or need to be reinterpreted in ways consistent with the idea of a nominal definition. One such way is to reinterpret the problem of the definition of emotion as a problem of terminological standardization, that is, of getting emotion researchers to agree to a stipulation about how to use the term “emotion” (Scherer, 2005). Another way is to reinterpret the question “What is an emotion?” as asking for the meaning of emotion words in ordinary language (e.g., “What do people mean by ‘fear’?”) (Hempel, 1965), with the implicit assumption being that the ordinary language definitions of emotions are themselves nominal (i.e., based on linguistic conventions).

However, since the demise of logical empiricism and the rise of the “new scientific realism” philosophies of science, the idea that some definitions—in fact, those of greatest interest to science—are after all a form of real definitions has gained renewed respectability (for details, see Boyd, 1991; 2002; Griffith, 1997). According to the modern view, a real definition is a hypothesis about the “deep structure” or core constitution of the objects in the class picked out, more or less precisely, by an extension-fixing device, such as a list of examples or a set of typical features. The paradigm case is the scientific definition of water, identified by a set of features such as “the clear, odorless liquid found in lakes and rivers, essential for animal and plant life (etc.)” as “H₂O.”

In line with others (e.g., Griffith, 1997; see also Charland, 2002), I believe that this “real definition” view is the correct model for the definition of emotions (Reisenzein, 1994a; 2007; Reisenzein & Schönplflug, 1992). According to this understanding, a definition of emotion is an empirical hypothesis about the nature or constitution of the states identified, more or less precisely, with the help of a list of paradigmatic examples (e.g., joy, sadness, fear, hope, disappointment, pity, joy for another...) and a set of typical features of the items on this list (a “working definition” of emotions; see Meyer, Reisenzein, & Schützwohl, 2001). This hypothesis about the nature of emotion is always formulated against the background of a theory of emotion generation (and the effects of emotions). Hence, the definition of emotions always presupposes such a causal theory, and it stands and falls with this theory (Reisenzein & Schönplflug, 1992; Reisenzein, 2007).

Accordingly, the title question of this article, “What is an emotion *in BDTE*?” asks for a theoretical identification of emotions while presupposing the truth of BDTE. This means at minimum that the *basic assumption* of BDTE—beliefs and desires are necessary for emotions—is accepted. However, when comparing Castelfranchi and Miceli’s (2009) view of the nature of emotion in BDTE with my own view, a richer version of BDTE can be presupposed that includes additional shared assumptions; in particular, the assumption that beliefs and desires cause feelings of pleasure and displeasure.

⁷ Supporting this view, the history of emotion psychology is full of attempts to empirically test proposed definitions of emotion (e.g., those of James, 1890; or Schachter, 1964). The view that an emotion definition is a nominal definition would make this research activity look utterly irrational. These empirical tests, however, make perfect sense if an emotion definition is a real definition—an empirical claim about the essence of emotions (Reisenzein, 1994a).

Given this understanding of the problem of the definition of emotion, how can this problem be solved? In principle, I propose, it can be solved by means of an inference to the best explanation (e.g., Lipton, 2004). That is, different proposals concerning the nature of emotion in BDTE—different proposals for the theoretical definition of emotions—are compared in terms of their ability to explain various accepted properties of emotions, such as their type distinctions (happiness, fear, etc.), their experiential quality, their intensity, their (apparent) object-directedness, and their temporal course; and the theoretical definition of emotions that best explains these properties of emotions is (provisionally) accepted.

One last point: Although I have discussed the question of the definition of emotion for emotions *in general*, it cannot be *directly* answered on this level. The reason is that emotions always come in more or less specific qualities (happiness, fear, etc.), and that the set of emotions demarcated by a working definition has fuzzy boundaries and may comprise several distinct subgroups (e.g., “cognitive” versus “sensory” emotions). What one can do, however, is to try to answer the definition question for as many paradigmatic emotions as possible and try to make plausible that parallel answers can be given in other similar cases (see also, Ortony et al., 1988). Furthermore, provided that the presupposed theory of emotion accounts well for paradigmatic emotions, one can and should use the theory itself to help decide on its range of application (Reisenzein & Schönplflug, 1992; Reisenzein, 2009b). This may lead to certain theoretically motivated reclassifications. For example, in CBDTE, surprise is classified as an emotion (Reisenzein, 2009a; 2009b).

In the remainder of this article, I will put the proposed method to work with the aim of deciding which of the two described theories of the nature of emotion in BDTE is more plausible. My conclusion will be that the gestalt theory of emotion is less successful overall in explaining salient properties of emotions than the causal feeling theory. My argumentation for this conclusion proceeds in three steps. In the first step, I present two arguments for the assumption, made by both Castelfranchi and Miceli and myself, that the feeling of pleasure or displeasure is a necessary component of emotions: I argue that this assumption is the best explanation of the phenomenal character of emotional experiences and their intensity. In the second step, I examine the two main arguments for regarding beliefs and desires as emotion components—that doing so is required to explain the finer distinctions among emotions and their object-directedness—and argue that they are unconvincing. In step three, I present three objections against the hypothesis that beliefs and desires are components of emotions. The causal feeling theory avoids these objections. I therefore recommend abandoning the belief-desire compound theory of the nature of emotion in favor of the causal feeling theory.

3. Why Hedonic Feelings (Pleasure and Displeasure) are Necessary for Emotions

Castelfranchi and Miceli (2009) and I agree on an “affect-enriched” version of BDTE: We assume that “emotional” belief-desire constellations cause feelings, most importantly of pleasure or displeasure, and that these feelings are essential for emotional experience. What we disagree on is the more precise theoretical definition of emotion, given these shared assumptions. I propose to identify emotions with the output signals of the emotion-generating mechanisms assumed in CBDTE (the BDC and the BBC) that, when conscious, are experienced as feelings of pleasure and displeasure, surprise and expectancy confirmation, mix-

tures of these, and hope and fear (Reisenzein, 2009a; 2009b). By contrast, Castelfranchi and Miceli (2009) propose to identify emotions—which they essentially equate with emotional experiences—with “mental compounds,” which they claim emerge from beliefs, desires, and hedonic feelings through a further information-processing step, a process that integrates beliefs, desires, and feelings into an emotional gestalt. Which of these theoretical definitions of emotion is more plausible?

As the first step in trying to answer this question, I will present arguments for the *shared assumption* of the two compared views of the nature of emotion: the assumption that feelings of mental pleasure or displeasure are (at least) a necessary component of (most) emotional experiences. Castelfranchi and Miceli (2009) do not argue for this assumption, probably because it was introspectively evident to them, as it was to numerous classical (e.g., Bentham, 1889/1970; Külpe, 1893; Meinong, 1906; Wundt, 1896) and contemporary emotion researchers (e.g., Barrett, 2006; Cabanac, 2002; Goldstein, 2002; Mellers, 2000; Ortony, Clore, & Collins, 1988; Reisenzein, 1994b; 2009b; Russell, 2003). As a matter of fact, however, not everybody is convinced that emotions contain a hedonic feeling component; therefore, it *is* necessary to present arguments for this claim. The strongest arguments are, I believe, that this assumption provides the best available explanation for (a) the phenomenal character of emotions and (b) their intensity. Note that, to the degree that these arguments are successful, they place an important constraint on theories of the nature of emotions: Any plausible theory of the nature of emotions would then have to assume that (most) emotions contain a hedonic feeling component.

3.1 Explaining the Phenomenal Character of Emotions

Probably the most salient property of emotions, from the perspective of the experiencing person, is their phenomenality, the fact that it “is like” or “feels” a particular way for the person to have an emotion (e.g., Reisenzein & Döring, 2009). Although there is a long-standing and still ongoing debate about how many different affective feeling qualities there are, both everyday experience and more formally collected psychological data suggest that, barring hedonically neutral surprise (which is included in CBDTE, for theoretical reasons, as a limiting case of emotions), all emotional experiences are at least among others characterized by a pleasurable or displeasurable quality, a positive or negative hedonic tone (see e.g., Barrett, Mesquita, Ochsner, & Gross, 2007; Külpe, 1893; Mellers, 2000; Reisenzein, 1994b; Russell, 2003; Reisenzein & Junge, 2006; Wundt, 1896). The phenomenal quality, specifically the hedonic tone of emotional experiences, thus needs to be accounted for by any plausible theory of emotion.

The assumption that “emotional” constellations of beliefs and desires cause feelings of pleasure or displeasure, and that these feelings are necessary components of emotional experiences, provides a natural explanation for the hedonic tone of emotions (Reisenzein, 2009b). A feeling of mental pleasure or pain, conceived of as a sensation-like mental state similar to sensations of color, tone, or temperature, explains the hedonic quality of emotions in a natural way because it belongs to the essence of sensations to have a phenomenal quality (Külpe, 1893; Wundt, 1896). In fact, I believe that the assumption that emotions contain—or even are—feelings of mental pleasure or mental pain provides the *best available* explanation of the hedonic tone of emotions. However, I will first argue for a weaker claim: To account for the hedonic tone of emotions in BDTE, one must assume that “emo-

tional” belief-desire configurations cause *a separate mental state that carries the hedonic tone*. The main argument for this conclusion is that attempts to explain the hedonic character of emotions in terms of belief and desire alone seem to meet with insuperable difficulties. To see this, let us consider how BDTE theorists might try to account for the hedonic character of emotions in terms of belief and desire (for additional discussions, see Pugmire, 1998; Salmela, 2002). In principle, these theorists have two options. First, they can try to explain the hedonic quality of emotions by appealing to the phenomenal properties of beliefs and desires. Second, they can assume that hedonic tone is an emergent property of the fusion of “emotional” belief-desire configurations (Green, 1992). Let us examine these two options in turn.

3.1.1 Explaining Hedonic Quality by Appealing to the Phenomenal Properties of Beliefs and Desires

The first stepping stone for BDTE theorists who try to explain the experiential quality of emotions by appealing to the phenomenal qualities of beliefs and desires is the intuition that these mental states do not have any experiential quality (e.g., Green, 1992; Smith, 1987; for a recent overview of this debate about “cognitive phenomenology,” see Bayne & Montague, 2011). According to this intuition, being conscious of our occurrent beliefs or desires consists exclusively of our being immediately (noninferentially) aware of them when they occur; but there is nothing it is like to have them.

However, even assuming that beliefs and desires do have phenomenal character, as some have argued (see Bayne & Montague, 2011; I concede this possibility in Reizenstein, 2009a), it is questionable whether this character is *of the right kind* to explain the experiential and specifically the hedonic quality of emotions.

(1) The phenomenal qualities of belief and desire, considered separately, seem unsuited to explain the hedonic tone of emotions. Beliefs—at least the factual beliefs regarded as preconditions of emotions in BDTE—do not have any intrinsic pleasure-pain quality. With respect to conscious desires, it is more plausible to argue that they are occasionally characterized by hedonic tone (e.g., Marks, 1982). However, often they are not, and when they are, they seem to be mainly unpleasant, particularly when they are intense (e.g., cravings). Furthermore, the hedonic tone of an emotion is sometimes the opposite of the hedonic tone most plausibly attributed to the underlying desire. The satisfaction of unpleasant cravings is still pleasurable; likewise, the hedonic tone of relief about the non-occurrence of an undesired state of affairs *p* is pleasurable, whereas the hedonic tone of the underlying desire (the aversion to *p*) is, if anything, negative.⁸ Finally, the pleasure of desire fulfillment cannot be the hedonic tone of the desire because desires are extinguished upon their subjective fulfillment (see Section 5.2).

(2) If these intuitions are correct, then the hedonic quality of emotions also cannot be explained in terms of a mixture of the experiential qualities of beliefs and desires—if only because the belief, lacking hedonic tone, cannot contribute anything useful to that mixture.

⁸ When reflecting on the hedonic tone of desires, one should keep in mind that one becomes aware of a desire mainly when one acquires the belief that a state of affairs that fulfills or frustrates the desire is possible or certain—that is, at the time when an emotion occurs. There is thus the danger of attributing the hedonic quality of emotions to the desire, in which case the desire account of the hedonic tone of emotions would be circular (see also Stocker, 1983).

3.1.2 Hedonic Quality as an Emergent Property

To overcome the difficulties of explaining the hedonic quality of emotions in terms of the phenomenal qualities of beliefs and desires, Green (1992) proposed that emotions are the products of a fusion of “emotional” belief-desire constellations, and that hedonic quality is an emergent property of the resulting mental states.⁹ If we assume, as I think we must, that the proposed fusion process is a causal process, Green’s suggestion combines two claims: (a) “Emotional” belief-desire constellations cause another mental state with hedonic properties, and (b) this mental state is a holistic compound of beliefs and desires. Of these two claims, the second remains largely metaphorical until the details of the proposed fusion process are spelled out. However, about all that is said about this process is that its inputs are beliefs and desires, whereas its output is a mental compound, which contains beliefs and desires as parts, and has emergent properties including hedonic quality and intensity.¹⁰ Hence, this second claim is at best a promissory note of future explanation.¹¹ Therefore, we are left with the first claim. However, the first claim is identical to the hypothesis I presently seek to establish: The carrier of hedonic tone is a mental state caused by beliefs and desires.

The discussed options seem to exhaust the possibilities of explaining the hedonic quality of emotions in terms of the beliefs and desires that, according to BDTE, are necessary conditions of emotions. The failure of these explanatory attempts implies that beliefs and desires alone are insufficient to explain the hedonic quality of emotions; an additional mental state caused by “emotional” belief-desire constellations needs to be assumed. This conclusion does not per se entail that this additional component of emotion, the carrier of hedonic tone, is a sensation-like, intrinsically objectless feeling, as Reizenzein (2009a; 2009b) and Castelfranchi and Miceli (2009) assume. There is at least one alternative possibility. Illustrated for the case of happiness about p , this is the proposal that the belief that p and the desire for p cause another intentional mental state directed at p , $F(p)$, that is *simultaneously* a feeling of pleasure. The theory that emotions are, or at least comprise, “intentional feelings” (specifically, of pleasure or displeasure) was originally advanced by Franz Brentano and his students, in particular, Alexius Meinong (1894; 1906) and Carl Stumpf (1899; see Reizenzein and Schönplflug, 1992) and has recently been repropounded by several philosophers (e.g., Goldie, 2000; Helm, 2001). Although attractive in other respects—in particular because it

⁹ Hence, Green assumes that hedonic tone is a property of the *output* of the fusion of “emotional” belief-desire constellations. By contrast, Castelfranchi and Miceli (2009) assume that feelings of pleasure or displeasure are one of several *inputs* to the fusion process that results in the emotion.

¹⁰ In addition, Green (1992, p. 97) draws an analogy with the process of chemical bonding (e.g., of hydrogen and oxygen into water). However, this analogy is presumably meant only to document the existence of physical compounding processes whose products have emergent properties, and thereby to raise the probability that analogous processes might exist in the mental realm; it is not meant to imply any deeper similarity of the intervening processes.

¹¹ Attempts to provide this explanation suggest that the proposed fusion process is actually something else entirely. Specifically, CBDTE suggests two things about this process: (a) It includes the comparison of newly acquired beliefs with existing beliefs and desires, for without the detection of a belief-desire match or mismatch, no emotion will result (Reizenzein, 2009a); (b) beyond the belief-desire comparison, no further process is needed to explain the hedonic quality, as well as the intensity, of emotions. Hence, CBDTE suggests that the process in question is one of comparing beliefs and desires, not of fusing them into a new mental state.

offers a neat explanation of the intentionality of emotions—the theory of intentional pleasure feelings has to cope with a number of difficulties: It has been debated whether propositional attitudes have phenomenal character (see Bayne & Montague, 2011), and it is not clear what the semantic properties (conditions of satisfaction; see Green, 1992) of object-directed feelings would be. The hypothesis that the carrier of the hedonic tone of emotions is an objectless sensation-like feeling of pleasure or displeasure avoids these problems. For this and other reasons (Reisenzein, 2009a; b), I prefer this hypothesis to the “intentional feeling” theory. However, it should be noted that the “intentional feeling” theory is similar enough to the “sensory feeling” theory to be regarded, at least in the present context, as a variant of the latter: Both theories assume that the affective component of emotions is a hedonic feeling, but in one case this feeling is intrinsically objectless, whereas in the other case, it is directed at the object of the emotion.

3.2 Explaining the Intensity of Emotions

Emotions are not just present or absent; rather, they can be instantiated in different degrees or gradations, ranging from just noticeable to extremely intense. For example, one can be a little, moderately, or very happy; or mildly, somewhat, or extremely surprised. Intensity has been neglected in theoretical discussions of emotion. Nevertheless, it is an undisputed, salient feature of emotions that, therefore, any plausible theory of emotion must explain (Frijda, Ortony, Sonnemans, & Clore, 1992; Green, 1992; Pugmire, 1998; Reisenzein, 1994b). I submit that, just like the phenomenal quality of emotions, the intensity of emotions cannot be explained in terms of belief and desire alone. To account for the intensity of emotions, another component of emotions needs to be posited; and the best candidate for this additional component, I submit, is a sensation-like feeling.

In CBDTE, this idea is implemented as follows: It is assumed that “emotional” constellations of beliefs and desires cause sensation-like feeling qualities, and that the intensity of these feelings is a quantitative function of belief and desire strength (Reisenzein, 2009a; 2009b). Because quality and intensity are essential properties of sensations (Külpe, 1893; Wundt, 1896), these assumptions allow CBDTE to simultaneously explain the phenomenal quality and the intensity of emotions in a natural way: Emotions are signals produced by the belief-desire and belief-belief comparator mechanisms that are experienced as feelings, and the intensity of the emotions is simply the intensity of these feelings (Reisenzein, 2009b). Again, although I believe that this is the best available explanation of emotion intensity, I will first argue for a weaker claim: To account for the intensity of emotions in BDTE, it is necessary to posit a separate mental state caused by beliefs and desires, one that carries emotional intensity.

One argument for this conclusion can be obtained by an extension of the preceding argument concerning the explanation of hedonic quality. The hedonic quality of emotions itself has an intensity; it can occur in different degrees (Reisenzein, 1994b). If the carrier of the hedonic tone of emotions is a separate mental state caused by beliefs and desires, then the different intensities of pleasure and displeasure must be carried by (different specifications of) this separate mental state.

However, the conclusion that the intensity of emotions is carried by a separate mental state can also be arrived at independently of the phenomenal quality argument, by considering the problems that confront BDTE theorists who attempt to explain the intensity of emo-

tions in terms of belief and desire alone. Again, these theorists have two principal options: They can try to explain the intensity of emotions by appealing to the intensities of beliefs and desires, or they can assume that emotion intensity is an emergent property of the fusion of “emotional” belief-desire constellations (Green, 1992).

3.2.1 Explaining Emotion Intensity by Appealing to the Intensities of Beliefs and Desires

(1) The intensities of belief and desire, considered separately, are unsuited to explain the intensity of emotions. Equating emotion intensity with belief strength does not work for the simple reason that belief strength has a clear upper bound (certainty), whereas the intensity of emotions has no definite upper bound. Furthermore, if emotion intensity were identified with belief strength, then emotions connected to beliefs held with certainty (the “emotions of certainty”; Green, 1992; Meinong, 1894), such as joy, could not vary in intensity. Neither can emotion intensity be equated with the strength of desire. In the case of the “emotions of certainty,” this idea may at first seem to have some plausibility: Mary feels more joy about Schroiber being elected chancellor the more she desires this event to happen; so perhaps the intensity of Mary’s happiness is simply the intensity of the desire. However, the intensity of the “emotions of uncertainty” (Green, 1992; Meinong, 1894), hope and fear, depends on *both* belief strength and desire strength. For example, Mary is more strongly afraid that she may miss her connecting plane the more probable this event appears to her and the more averse she is to it. Analogously, the intensities of disappointment and relief depend on both belief and desire strength (for empirical evidence, see e.g., Reizenzein & Junge, 2006).

(2) The intensity of emotion also cannot be equated with the intensity of “emotional” belief-desire configurations. The intensity of these configurations is, minimally, a two-dimensional magnitude of the kind $\langle b(p), d(p) \rangle$, whereas the intensity of emotion is a one-dimensional quantity, and is at least in some cases (e.g., hope, fear, disappointment, and relief) a joint function of the intensities of the involved beliefs and desires (Reizenzein, 2009a).

3.2.2 Emotion Intensity as an Emergent Property

To overcome these difficulties, one could again propose that emotions result from a fusion of “emotional” belief-desire constellations, and that the intensity of emotions is an emergent property of the resulting mental states (Green, 1992). However, this proposal provokes objections exactly parallel to those raised against the “emergentist” explanation of hedonic quality. One part of the proposal—“emotional” belief-desire constellations cause another mental state that is the carrier of emotion intensity—is identical to the hypothesis I presently seek to establish, and thus grants this hypothesis. The second part of the proposal claims that the new mental state, the carrier of emotion intensity, is a compound of beliefs and desires. This claim is at best a promissory note for an explanation as long as the fusion process is not spelled out, and CBDTE suggests that this process is actually one of comparing, rather than fusing, beliefs and desires.

The failure of the attempts to explain the intensity of emotions in terms of the intensities of belief and desire suggests that the latter are insufficient to explain the intensity of emo-

tions. To explain emotion intensity in BDTE, an additional mental state, caused by “emotional” belief-desire constellations, needs to be posited. This conclusion does not per se entail that this mental state, the carrier of emotion intensity, is a sensation-like, intrinsically objectless feeling. An alternative is again that emotions are a class of mental states *sui generis*, which are simultaneously feelings *and* object-directed. As mentioned in Section 3.1.2, the “intentional feeling” hypothesis has a number of problems that make me prefer the sensory feeling hypothesis. However, in the present context, the “intentional feeling” and the “sensory feeling” hypotheses of the nature of emotions can be regarded as variants of one and the same proposal.

4. Two Unconvincing Arguments for the Belief-Desire Compound Theory of Emotion

Because phenomenal quality and intensity are essential properties of sensations, the assumption that emotions contain a sensory feeling component makes it possible to explain both the phenomenal quality and the intensity of emotional experiences in a natural way. Why then not simply assume that emotions *are* the feelings of pleasure and displeasure (and a few others, such as surprise) caused by the detection of, inter alia, desire-fulfillment and desire-frustration, as assumed by CBDTE? Why assume that beliefs and desires, too, are *components* of the emotion, as Castelfranchi and Miceli (2009) propose? Two main arguments have been advanced for this theory of the nature of emotions: the *emotion differentiation* argument and the *intentionality* argument. It has been claimed that only by conceiving of beliefs and desires as parts of the emotion can one explain (a) the finer distinctions between emotions—possibly all distinctions beyond positive and negative—and (b) the intentionality or object-directedness of emotions. However, although these arguments have been influential, they are ultimately unconvincing (see also Reisenzein, 2000; Reisenzein & Döring, 2009).

4.1 Why the Emotion Differentiation Argument Fails

Castelfranchi and Miceli (2009) refer to the emotion differentiation argument to motivate their theory that emotions are compounds of beliefs, desires, and affect. They argue that, although “the feeling of pleasantness and unpleasantness [is]...one essential component of emotional experience”, beliefs and desires need to be added “to produce distinct emotions” (pp. 223-224). However, one can accept that beliefs and desires are necessary to distinguish between emotions without assuming that emotions contain beliefs and desires as (ontological) parts. A main reason for this is that mental states need not—and, if causal-role functionalism about mental states (e.g., Block, 1980) is right, cannot—be defined and distinguished from one another exclusively in terms of their intrinsic or nonrelational properties; they can also be defined and distinguished in terms of their relational features, in particular, their causes and consequences. Partly causal definitions are in fact common in everyday language. To mention a standard example, “sunburn” is defined as an “inflammation of the skin caused by over-exposure to sunlight” (Gordon, 1978, p. 125). Analogously, emotions could be defined as the feelings (e.g., of pleasure or displeasure) that are caused by particu-

lar constellations of beliefs and desires (see also Reisenzein, 1994a; 1994b; Reisenzein & Schönplflug, 1992). For example, happiness about p could be defined as the feeling of pleasure caused by the desire for p and the belief that p is the case. In fact, this is how emotional experiences are defined in CBDTE (Reisenzein, 2009a).¹² In adopting a (partly) causal definition of emotions, CBDTE finds itself in agreement with the position of causal-role functionalism in the philosophy of mind, the metaphysical backbone of contemporary cognitivism (e.g., Block, 1980; Fodor, 1975). It also finds itself in agreement with the views of many cognitive emotion theorists from Aristotle (about 350 B.C.) to Arnold (1960)¹³ and with the folk-psychological understanding of mentalistic terms in general (Lewis, 1972) and emotion terms in particular (e.g., Reisenzein & Junge, 2012; Siemer, 2008).

The existence of causal-functional definitions of emotions (in both scientific and common-sense psychology) means that the emotion differentiation argument is invalid: The claim that beliefs and desires are necessary to distinguish between emotions does not entail that beliefs and desires are components of emotions. The belief that this inference is logically valid may have been due to an inadvertent confusion of the conceptual-linguistic and ontological levels of analysis: the *concepts* “belief,” “desire,” “happiness,” “fear,” etc., and the *referents of these concepts*. This confusion could have been facilitated by the fact that part-whole relations exist on both levels. Suppose that an emotion researcher sets out to clarify the nature of emotions and presents the results of this analysis in the form of a theoretical definition that states necessary features. To illustrate, suppose that the researcher defines or analyzes the concept “being happy about p ” using the concepts “believing p ” and “desiring p .” In this case, the latter concepts are components of the former concept in the sense that they (or the terms that express them) are parts of the proposed definiens or analyzans of “being happy about p .” It might then be concluded that the *referents* of “believing p ” and “desiring p ” (the mental states of believing and desiring) are likewise parts of the *referent* of “being happy about p ” (the mental state of happiness). However, this would be too quick: The concept “being happy about p ” could be defined functionally, for example, as “the feeling of pleasure caused by the belief that p and the desire for p .”

Some proponents of the thesis that emotions contain beliefs and desires (or other mental states) as parts may in fact only have had a conceptual part-whole relation in mind, with no implications for the ontological level. In previous writings on emotion, Miceli and Castelfranchi (e.g., 2002), too, seem to use the expression “belief B and desire D are compo-

¹² To be precise, emotions and emotional experiences are defined in CBDTE on the computational level of system analysis. *Emotions* are the output signals of hardwired mechanisms that compare new beliefs with existing beliefs and desires (Reisenzein, 2009a, b). These signals can remain unconscious (e.g., when they are below a minimum level of intensity), but usually, they give rise to *emotional experiences*, which are conscious feelings of pleasure and displeasure, surprise and expectancy-confirmation, mixtures of these, plus hope and fear. On the phenomenological or intentional level of system analysis—that knows nothing about the representational codes and computational mechanisms underlying beliefs and desires—these experiences can be defined as emotional feelings that are caused by particular belief-desire configurations.

¹³ In his *Rhetoric*, Aristotle defined fear as a kind of displeasure or perturbation arising from the idea of impending evil. Descartes (1649) defined several of his nonbasic emotions as subtypes of basic feelings whose distinguishing features consist of their being caused by particular types of appraisals. Stumpf (1899) proposed that emotions are belief-caused pro- or con-evaluations of states of affairs (see Reisenzein & Schönplflug, 1992). Arnold (1960) defined emotions as felt action tendencies caused by cognitive appraisals. In previous work, I proposed a causalist version of appraisal theory, according to which emotions are appraisal-caused mixtures of pleasure or displeasure and activation or deactivation (Reisenzein, 1994b). Further examples of functional definitions of emotion are referenced in Reisenzein and Schönplflug (1992).

nents of emotion E” in the conceptual part-whole sense only. Understood in this way, the claim that beliefs and desires are parts of emotions does not go beyond the basic claim of BDTE, that beliefs and desires are necessary for emotions, and is thus compatible with different ontological interpretations. However, in their 2009 article, Castelfranchi and Miceli take a stronger stance on the issue: There, they assume that beliefs and desires are ontological parts of the emotion.

In taking this stronger position, Castelfranchi and Miceli (2009) could have been motivated by the consideration that, even though the assumption that beliefs and desires are ontological parts of emotions may not be needed to explain the differences between emotions *in general*, it may be needed to explain their differences in *phenomenal quality*. That is, it could be argued that conceiving of beliefs and desires as mere causes of emotions does not explain how different emotions can *feel* different to the experiencer in ways that go beyond pleasure and displeasure (and perhaps a few other feelings). To explain, for example, why happiness *feels* different from pride, one must appeal to the components of emotional experiences, rather than to their causes.

However, this argument is also unconvincing. For one reason, its implicit general premise—nothing short of a proper part of an experience can contribute to that experience’s phenomenal character—can be doubted. The mental context of an emotional feeling, including its causes and consequences, might conceivably change the experience of that feeling (Reisenzein, 2009b). For another reason, even if one accepts the general premise of the argument, the conclusion does not follow. First, there may in fact be nothing to explain: The argument may rest on a confusion of the differences between *emotions* with differences in the *phenomenal quality of emotions*. There is no question that, for example, happiness and pride differ—they differ, in particular, in their cognitive-motivational preconditions. However, this does not necessarily mean that happiness and pride *feel* different: Not every difference between emotions needs to be a felt difference, a difference in experiential quality. In terms of experiential quality, happiness and pride might be exactly alike: both feel pleasant (e.g., Bentham, 1789/1970; Külpe, 1893); any cognized differences between happiness and pride could be due to differences in the perceived causes (and consequences) of the feeling of pleasure. Second, regarding beliefs and desires as emotion components may not be of much help in explaining the phenomenal quality of emotions, for as mentioned before (3.1.1), beliefs and desires may not have phenomenal properties, or at least none that could explain differences in emotional experience. In this context, it is important to note that the subtler distinctions between “emotional” belief-desire constellations concern differences in their *contents*—differences in *what* is believed and desired (see Castelfranchi and Miceli, 2009, for examples). To explain presumed subtle experiential differences between emotions in terms of belief and desire, one must therefore assume that at least some differences in the contents of beliefs and desires are reflected in phenomenal experience (see Bayne & Montague, 2011).

4.2 Why the Intentionality Argument Fails

The second introspectively salient property of emotional experiences, apart from their phenomenal quality (including their intensity) is their (apparent) intentionality or object-directedness: Typically at least, emotions present themselves to the experiencer as being directed at certain objects. For the emotions that fall within the purview of BDTE, these ob-

jects are propositions or states of affairs (e.g., Mary feels happy about the fact that Schroiber was elected chancellor). The intentionality argument holds that, to account for the object-directedness of emotions, one must assume that beliefs and desires are not just causes but components of emotions (see e.g., Green, 1992; Pitcher, 1965; Reizenzein & Schönflug, 1992; Solomon, 1976). The intentionality argument is arguably the strongest argument for regarding cognitions (as well as, in the case of BDTE, desires) as components of emotions (Green, 1992; see also, Whiting, 2011). Nevertheless, this argument is not water-proof either. There are at least two ways in which the causal feeling theorist can avoid the conclusion of the intentionality argument.

The first way out is to accept the initial premise of the argument, that emotions are object-directed, but argue that emotions are a separate kind of object-directed mental states, distinct from but caused by beliefs and desires. In Sections 3.1.2 and 3.2.2, I mentioned a special version of this idea: the hypothesis that the feeling of pleasure caused by the belief that p and the desire for p is intrinsically directed at p .¹⁴ However, as also mentioned there, it is debatable whether propositional attitudes can have phenomenal quality, and it is unclear what the semantic properties of object-directed feelings would be. For these reasons, I prefer the second option available to the causal feeling theorist; one that avoids these problems. This option consists of denying the initial premise of the intentionality argument, that emotions are object-directed. Note that this is not meant to deny that emotions *appear to be*, in a presystematic sense of the term, “directed at” objects; for example, that Mary’s happiness appears to her to “focus” on Schroiber’s election victory. Rather, the argument is that this subjective appearance of focus does not reflect a genuine intentional relation (Reizenzein, 2009a). That is, whereas the experienced object focus of emotions may be *reminiscent* of true intentional (i.e., representational) relations such as believing p or desiring p , closer examination reveals that it is not a genuine representational relation after all.

The plausibility of this suggestion depends, among others, on whether a plausible explanation of the illusion of intentionality of emotions can be given. CBDTE offers an explanation: According to this theory, the illusion of object-directedness can be traced to the special way the emotion-generating mechanisms operate (Reizenzein, 2009a). For example, when the belief-desire comparator detects that the contents of a newly acquired belief $Bel(p)$ match the contents of an existing desire $Des(p)$, it generates a feeling of pleasure and, simultaneously, focuses attention on the responsible proposition p . It then appears to the person that she is *pleased about* p . Furthermore, it seems conceivable that subsequent cognitive processes bind feelings to the representations of the objects of the beliefs that proximately caused them, transforming these initially neutral thoughts into “affectively tinged” thoughts about these objects (Reizenzein, 2009a; 2009b; see also James, 1890/1950). For example, the feeling of pleasure caused by Mary’s belief that Schroiber won the election might get attached to the mental representation of Schroiber’s election victory, resulting in Mary’s thought of a pleasurable victory. Alternatively, the appearance of object-directedness might be due to an implicit causal attribution of the feeling (Reizenzein, 1994a; Schachter, 1964).

¹⁴ The gestalt theory of emotion, too, implies that emotions (emotional gestalts) are caused by beliefs and desires (Castelfranchi & Miceli, 2009). However, the resultant emotional gestalts are still thought to contain beliefs and desires as components. Therefore, this proposal is not sufficiently different from the part-whole theory to be regarded as a real alternative.

I conclude that the two main arguments in favor of the idea that beliefs and desires are components of emotions—the emotion differentiation argument and the intentionality argument—are unconvincing.

5. Three Problems for the Belief-Desire Compound Theory of Emotion

I argued that the two main arguments for the belief-desire compound theory of the nature of emotions are not compelling. However, this does not necessarily mean that this theory is wrong and the causal feeling theory is correct. It only means that, to decide this issue, we must turn to other arguments. In the following section, I present three arguments that speak directly against the belief-desire compound theory of emotion.¹⁵

5.1 The Problem of Emotional Intentionality – Still Unsolved

In Section 4.2, I argued that one need not regard beliefs and desires as ontological parts of emotions to explain the (apparent) intentionality of emotions: One can either deny that emotions are truly intentional (representational), in which case their object-directedness need not be explained; or one can argue that emotions are a separate class of intentional mental states caused by beliefs and desires. However, none of this touches the claim of the belief-desire compound theorists, that taking beliefs and desires to be components of emotions provides *another* viable explanation for the intentionality of emotions. Furthermore, belief-desire compound theorists could argue that their explanation is preferable because it neither requires questioning the intentionality of emotions nor introducing new questionable mental entities (object-directed feelings). I will now argue that the belief-desire compound theory of emotion, at least in its original form, is in fact *unable* to explain the intentionality of emotions. The reason is that this theory does not allow for correctly identifying the objects of emotions—“correctly” here meaning “the way we intuitively identify them in everyday life.”

Suppose Oscar is sad that he was not invited to a party. According to the belief-desire compound theory of emotion, Oscar’s sadness is a compound mental state that emerges, via a gestalt formation process, from Oscar’s belief that he was not invited to the party, and his desire to be invited (plus, in Castelfranchi and Miceli’s version of the theory, a feeling of displeasure). The object of Oscar’s sadness is the state of affairs described, from the first-person perspective, by the sentence “I am *not* invited to the party” (*not-p*), which is identical to the object of Oscar’s belief. The object of Oscar’s desire, however, is the state of affairs described by “I am invited to the party” (*p*). Hence, if *Bel(not-p)* and *Des(p)* are both components of Oscar’s sadness, it would seem that the emotion has *two* different objects that in addition, are contradictory opposites: *p* and *not-p*. But although Oscar is sad that he was not invited to the party (*not-p*), it is not the case that he is sad that he was invited (*p*); nor is it the case that he is sad about a contradiction (*not-p & p*). In fact, since Oscar does

¹⁵ The first two arguments also speak against a categorization account (e.g., Barrett, 2006; Mandler, 1984) of the “fusion” of beliefs and desires into an emotion.

not believe p , he cannot—according to BDTE—be sad about p . And if he believed p , he would be happy rather than sad about p .

Parallel cases can be constructed for other emotions including surprise—in general, for all emotions that occur when a desire is frustrated by a newly acquired belief, or when a previous belief is disconfirmed by a new belief. Take surprise: If Oscar expects to be invited to the party (p), and then comes to believe that he was not invited (not- p), he will be surprised about not- p . If belief-desire compound theorists of emotion analyze surprise in a way that parallels their analysis of hedonic emotions—that is, by including the necessary cognitive preconditions of surprise into the emotional compound—they have to regard both the belief that not- p and the belief that p as components of surprise. However, it is not the case that Oscar is surprised that he was invited to the party (p), nor is it the case that he is surprised about a contradiction (not- p & p). Likewise, if Oscar both desires and expects to be invited to the party (p), and then learns that he is not invited (not- p), he will be disappointed about not- p ; he will not be disappointed about p , nor about not- p & p .

In sum, emotions on the one hand, and their cognitive and motivational preconditions on the other hand, have the same propositional objects only in part. The object of the emotion corresponds to that of the proximate (according to CBDTE, the most recently acquired) emotion-relevant belief, such as in the present example, Oscar's belief that he is not invited to the party. By contrast, the propositional objects of the emotion-relevant desires and pre-existing beliefs can be contradictory opposites of the object of the emotion (e.g., Oscar's desire to be invited to the party, and his belief that he would be). Therefore, it seems that emotions cannot be identified with compounds that include the cognitive and motivational preconditions of the emotions—at least not all of them.¹⁶

To solve this problem, Castelfranchi and Miceli (2009) might appeal to the proposed gestalt formation process: They could argue that one effect of this process consists precisely of providing the emotional gestalt with a single appropriate object. However, without further elaboration, this escape is too simple. Similar to the attempt to explain the hedonic quality and intensity of emotions by postulating that emotions are the outcomes of a process of a fusion of beliefs and desires whose details are left unspecified (see 3.1.2 and 3.2.2), the gestalt formation process here plays the role of a *deus ex machina*—it is simply ascribed whatever causal powers are needed to yield an output with desired properties (in the present case, a mental state with a particular object). We still do not understand *why* the emotion has only one object and why that object is identical to the object of the “proximate” belief, rather than to the object of the desire, or that of a pre-existing belief. Shall we assume that the desire and the pre-existing belief lose their objects when they become part of the emotional gestalt? But in this case, they would cease to exist; it belongs to the essence of beliefs and desires to be directed at objects. Furthermore, if the object of the emotional gestalt is

¹⁶ Maria Miceli (personal communication) suggested that the object of Oscar's sadness might in fact not be not- p , but q = “my desire for p was frustrated.” However, intuitively this seems wrong: Oscar is first and foremost sad that he was not invited to the party (not- p); he is not, or at least not only, sad that his desire for p was frustrated. Analogously, Oscar is surprised that he was not invited to the party; he is not, or at least not only, surprised that his belief that p would occur was disconfirmed (Reisenzein, 2009a). Independent of this issue, because q is different from *both* the object of the belief (not- p) and the object of the desire (p) underlying Oscar's sadness, Miceli's proposal makes it even more difficult for belief-desire compound theorists to explain the specific object-directedness of the emotion. And if they assume that Oscar is sad about q because he desired not- q and then came to believe that q , the original problem arises again: Oscar is sad only about q ; he is not sad about not- q , nor about q & not- q .

the object of its belief component, can one really say that the *emotion*, rather than just its belief component, is directed at this object?

Even assuming that the problem of providing the emotional compound with the appropriate intentional object can be solved, another problem remains for Castelfranchi and Miceli's (2009) gestalt theory. Recall that according to this theory, emotions are mental gestalts that integrate beliefs, desires, and *hedonic feelings* (pleasure or displeasure). These hedonic feelings are regarded as intrinsically nonintentional. This raises the following question: Do the feelings become object-directed in the process of being integrated into the emotion? If yes, how is this feat achieved (Green, 1992)? If no, can one truly say that the *emotion* is directed at an object?

The more natural conclusion to draw from Oscar's case, it seems to me, is that at least desires and pre-existing beliefs are *not* parts of the emotion, but only their causes. Accepting this conclusion would mean, for Castelfranchi and Miceli's (2009) theory, that only the proximate belief and affect remain as components of the emotion (with the belief being a partial cause of the affect). This modified version of the belief-desire compound theory of emotion is already fairly close to the causal feeling theory. The main remaining difference concerns the question of whether the *proximate belief* is part of the emotion or not. I come back to this question at the end of the article.

5.2 Problems with Emotion Duration

A necessary requirement for the correctness of the belief-desire compound theory of emotion is that the elements of the compound are present during the emotion while it lasts. This is clear for the part-whole version of the belief-desire theory proposed by Marks (1982), where the emotion is *defined* as the co-existence of appropriate beliefs and desires. However, it is also the case for the gestalt theory version of BDTE (Castelfranchi & Miceli, 2009; see also Green, 1992, p. 81), for the gestalt-forming process produces an emotional gestalt only as long as it receives adequate inputs, the emotion components. To use Castelfranchi and Miceli's example of perceptual gestalt formation, the perception of a triangle exists only as long as the lines representing the sides of the triangle are perceived. If one of the lines is no longer visible, the perception of the triangle disintegrates.

I submit that this implication of the belief-desire compound theory of emotion—emotions are present only while their components are present—is not in line with the empirical facts. The discrepancy between theory and data is perhaps most apparent for desire-fulfillment emotions such as joy, but parallel discrepancies also exist for surprise and other emotional reactions to belief disconfirmation (e.g., disappointment, relief). In the case of the desire-fulfillment emotions, the problem is that, due to the updating of the belief-desire system that follows the detection of desire-fulfillment, the desire is usually deleted while the emotion still lasts. In the case of the expectancy-disconfirmation emotions, the problem is that the pre-existing belief with which a newly acquired belief is compared is changed while the emotion still persists. Another way of phrasing the problem is therefore that the temporal course, specifically the duration, of some emotions is not well explained by the belief-desire compound theory.

According to BDTE, Mary experiences happiness about p if she desires p and comes to believe that p is the case. Computationally speaking, happiness occurs if a cognitive subsystem (in CBDTE, the belief-desire comparator) detects that a desire has been fulfilled.

However, according to the usual understanding of “desire”, desires are extinguished by their detected fulfillment. As Meinong (1917, p. 96) put it, as soon as one comes to believe that a desired state of affairs exists, “the desire is destroyed; this is the subjective aspect of what is called...the fulfillment of the desire” (my translation). But if the desire for p is extinguished by the detection of its fulfillment, then the belief that p and desire for p are co-present in the cognitive system only as long as it takes to detect their congruence and to delete the desire (Reisenzein, 2009a). These are presumably primitive operations of the belief-desire updating mechanism that should accordingly take only minimal time (perhaps 300 ms; see Madl, Baars, & Franklin, 2011). Furthermore, strictly speaking, the emotion comes into existence only *after* desire-congruence has been detected. As a consequence, the emotion of joy should in general be an extremely short-lived occurrence, perhaps too short to be subjectively noticeable, rather than lasting for a while, as it typically seems to do (e.g., Fridja et al., 1992).

A parallel objection can be raised in the case of surprise. Surprise is experienced if a cognitive module (in CBDTE, the belief-belief comparator) detects that a newly acquired belief is contrary to an existing belief. For example, Mary expects Schroiber will not win the election, but then learns that he, in fact, won. As soon as the belief-belief comparator detects that the newly acquired belief contradicts the old, the old belief is deleted; a process that presumably takes only a fraction of a second. By contrast, according to subjective tracings of the temporal course of surprise, the experience of surprise typically lasts for several seconds (Reisenzein, Bördgen, Holtbernd, & Matz, 2006). And in the case of pleasant surprise, both the pre-existing desire and the pre-existing belief are updated while the emotion still persists.

Belief-desire compound theorists of emotion could try to solve this problem, in the case of the desire-fulfillment emotions, by bringing in other desires, such as, in our example, Mary’s desire that Schroiber will remain chancellor for as long as possible or that he will keep his campaign promises. But apart from the fact that it is not plausible that one experiences joy only if, at the subjective fulfillment of a desire, other related desires spring into existence, the problem is thereby not solved but only deferred. The desire that Schroiber will keep his campaign promises is different from the desire that Schroiber will win the election; correspondingly, the joy experienced when learning that Schroiber has kept his promises is different from the joy about Schroiber’s election victory. And the former feeling of joy should be experienced only when the corresponding wish is subjectively fulfilled and as a consequence, extinguished. Analogous objections can be raised against the attempt to explain the duration of surprise by bringing in additional disconfirmed beliefs.

A more promising strategy of dealing with the present objection would be to argue that, notwithstanding the above-mentioned theoretical considerations, fulfilled desires are in fact not extinguished immediately, and disconfirmed beliefs are not immediately deleted, but that the process of desire and belief updating takes time—just enough time for joy or surprise to occur and subside again. Emotions are experienced only while the desire and belief, or the old and new belief, are simultaneously present in the cognitive system, as assumed by the belief-desire compound theory. This seems to be an empirically testable difference between the belief-desire compound theory and the causal feeling theory; however, I do not know of firm empirical data that would allow for a decision regarding this issue. In the absence of supporting data, these auxiliary assumptions about the relative duration of belief and desire updates, and of the associated emotions, strike me as *ad hoc*. In any case, it should be realized that the belief-desire compound theory entails strong empirical assumptions about the relative duration of emotions and belief-desire updates. The causal feeling

theory does not require making these assumptions; it is equally well compatible with a variety of empirical data concerning this issue.

5.3 Are There Emotional Gestalts?

Castelfranchi and Miceli's (2009) gestalt theory of emotion would be on firmer ground if there were independent evidence that the emotional gestalts with which they identify emotions do exist (although they would then still have to argue for the *identity* of emotions and emotional gestalts). However, the existence of emotional gestalts is by no means certain. Emotional gestalts could be theoretical constructs that lack empirical referents. If so, the belief-desire compound theory of the nature of emotion cannot be correct, for surely, emotions cannot be identified with nonexistents (Reisenzein, 2007).

Whereas Castelfranchi and Miceli (2009) appeal to introspection to support the existence of emotional gestalts, I must admit that I do not share their intuitions. According to my introspection, when I am happy about some state of affairs p that fulfills a wish of mine, I am not experiencing a complex mental state that comprises the desire for p and the belief that p . Rather, I become aware that p is the case, and this thought is immediately followed and accompanied by an upsurge of pleasant feeling, that after a shorter or longer while, subsides again. I would also be willing to admit to an impression of phenomenal causality (Heider, 1958), in that the feeling appears to be caused by the thought that p . These introspective intuitions agree with those of others who have considered the issue (e.g., Meinong, 1906; Whiting, 2011). However, I am not aware of a fusion of belief and desire taking place (and as argued in 5.2, such an integration *cannot* take place because the desire is extinguished by its fulfillment). To the degree that anything like an integration occurs at all, it seems to concern only the object of the emotion p and the feeling: When thinking of the state of affairs p that I believe to be the case, it presents itself to me not just as a state of affairs that is real, but also as a pleasurable state of affairs (cf. James, 1890/1950).

These introspective worries about the existence of emotional gestalts of the kind required by Castelfranchi and Miceli's (2009) theory are reinforced by theoretically motivated concerns about the proposed gestalt formation process. As mentioned, the claim that emotions are the outcome of a process that integrates different components into a whole with suitable emergent properties runs the risk of being a pseudo-explanation—one simply imputes to the fusion process whatever causal powers are needed to produce outputs with the desired properties (e.g., intensity, a specific object). To avoid this danger, one must become more specific about the integration process. Castelfranchi and Miceli (2009) go some way toward explicating the integration process by proposing that it is analogous to gestalt-forming processes known from perception (e.g., Köhler, 1947). However, it should be acknowledged that the inputs of the emotional gestalt formation process (beliefs, desires, and hedonic feelings) differ significantly from the typical inputs of the perceptual gestalt formation processes (e.g., visual perceptions of lines), and so do their outputs (emotions versus e.g., the perception of geometrical figures). Given these peculiarities of the gestalt formation process in the case of emotions, it would be helpful to know that analogous gestalt formations, involving similar mental elements, occur in other, less controversial domains. However, gestalt formation processes have been nearly exclusively studied in the domain of perception. It therefore remains at present uncertain whether their extension to emotions really works.

6. The Causal Feeling Theory Vindicated

To recapitulate, I first argued that sensation-like feelings are a necessary component of emotions: This assumption is an inference to the best explanation of the phenomenal character (specifically the hedonic quality) of emotions and their intensity. Second, I argued that the two main arguments for regarding beliefs and desires as components of emotions—the emotion differentiation argument and the intentionality argument—are unconvincing. Third, I presented three objections to the hypothesis that beliefs and desires are components of emotions: This hypothesis fails, at second sight, to explain the directedness of emotions at specific objects; it has difficulty accounting for the duration of emotional reactions to the fulfillment of desires and the disconfirmation of beliefs; and there are reasons to question the existence of the postulated emotional gestalts and the process that presumably generates them.

The causal feeling theory of the nature of emotions avoids these problems. By assuming that emotional experiences are nonpropositional signals that, when conscious, are experienced as feelings, this theory accounts in a natural way for the phenomenal character of emotions and their intensity. It also vindicates our everyday talk of beliefs and desires as causes of emotions—something the belief-desire compound theory of emotion achieves only with some difficulty (see Footnote 4). The finer distinctions among emotions—those not attributable to differences in feelings—can be explained by assuming that emotions are, in part, distinguished by their cognitive and motivational causes (and consequences).

The price one may have to pay for these explanatory virtues of the causal feeling theory is the concession that—counter to linguistic practices and first phenomenological intuitions—emotions are *not* modes of representing propositional objects like beliefs and desires are. However, this may just be how things are. Furthermore, as mentioned, the belief-desire compound theory is incapable of explaining the specific object-directedness of emotions; it therefore presents no real advantage to the causal feeling theory in this respect. And by assuming that emotional feelings can be bound to the objects of the beliefs that cause them, resulting in “emotionally tinged” thoughts about them—or even more simply, that people form beliefs about the causes of their emotional feelings (Schachter, 1964; Reizenzein, 1994a; 1994b)—the cognitive (information-providing) and motivational functions of emotions can be saved without assuming that emotions themselves have propositional objects (Reizenzein, 2009b).

The last-mentioned elaboration of the causal feeling theory points to a possible reconciliation with the belief-desire compound theory: To solve the problems with explaining the intentionality and duration of emotions, and to avoid the uncertainties surrounding the postulated gestalt formation process and its outcomes, Castelfranchi and Miceli could modify their theory of the nature of emotions by proposing that only hedonic feelings and their proximate cognitive causes (beliefs), but not pre-existing beliefs and desires, are components of the emotion; the latter are only its causes. According to this revised version of the gestalt theory of emotion, Mary’s happiness about p would be a mental state that emerges from the integration of the belief that p , with the pleasure caused by this belief plus Mary’s desire for p . In the next step, the integration process could be explicated as the binding of sensory and propositional representations into an “affectively tinged” thought (Reizenzein, 2009a). The remaining difference to the causal feeling theory would then concern the question of whether the emotion is to be identified with the (signal underlying the) affective feeling, or with the “affectively tinged” thought resulting from the binding of the feeling to

the representation of p (in CBDTE, the sentence s in the language of thought representing p). I continue to believe that the former theoretical definition of emotion is more adequate because it accounts better, overall, for the different properties of emotions. In particular, the signal underlying the emotional feelings is the common cause of both emotional experience and the physiological and expressive effects of emotions (Reisenzein, 2009a; 2009b). However, an in-depth discussion of this question must be left to another occasion.

References

- Aristotle. (1980). *Rhetorik [Rhetoric]*. München: Fink. (Original work published about 350 b. c.).
- Arnold, M. B. (1960a). *Emotion and personality* (Vols. 1 & 2). New York: Columbia University Press.
- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*, 20-46.
- Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology*, *58*, 373-403.
- Bayne, T., & Montague, M. (2011, in press). Cognitive phenomenology: An introduction. In T. Bayne & M. Montague (eds.) *Cognitive Phenomenology*. Oxford: Oxford University Press.
- Bentham, J. (1789/1970). *An introduction to the principles of morals and legislation*. London: Athlone Press.
- Block, N. (1980). Introduction: What is functionalism? In N. Block (Ed.), *Readings in philosophy of psychology* (Vol. 1, pp. 171-184). Cambridge, MA: Harvard University Press.
- Boyd, R. (1991). Realism, Anti-Foundationalism, and the enthusiasm for natural kinds. *Philosophical Studies*, *61*, 127-48.
- Boyd, R. (2002). Scientific realism. *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2011/entries/scientific-realism/>
- Bratman, M. E. (1987). *Intentions, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Cabanac, M. (2002). What is emotion? *Behavioural Processes*, *60*, 69-84.
- Castelfranchi, C. & Miceli, M. (2009). The cognitive-motivational compound of emotional experience. *Emotion Review*, *1*, 223-231.
- Charland, L. (2002). The natural kind status of emotion. *British Journal for the Philosophy of Science*, *53*, 511-537.
- Conte, R., & Castelfranchi, C. (1995). *Cognitive and social action*. London: University College London Press.
- Davis, W. (1981). A theory of happiness. *Philosophical Studies*, *39*, 305-317.
- Descartes, R. (1984). *Les passions de l'ame [The passions of the soul]*. Hamburg, Germany: Meiner. (Original work published 1649)
- Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences* (pp. 572-595). Oxford: Oxford University Press.
- Feather, N. T. (Ed.). (1982). *Expectations and actions: Expectancy-value models in psychology*. Hillsdale, NJ: Erlbaum.
- Fodor, J. A. (1975). *The language of thought*. New York: Crowell.
- Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
- Fox, C. R., & Poldrack, R. A. (2009). Prospect theory and the brain. In Glimcher, P. W., Camerer, C. F., Fehr, E., & Poldrack, R. A. (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 145-174). London: Elsevier/Academic Press.
- Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Frijda, N. H., Ortony, A., Sonnemans, J., & Clore, G. L. (1992). The complexity of intensity: Issues concerning the structure of emotion intensity. In M. S. Clark (Ed.), *Review of Personality and Social Psychology* (Vol. 13, pp. 60-89). Newbury Park: Sage.
- Goldie, P. (2000). *The emotions: A philosophical exploration*. Oxford: Clarendon Press.
- Goldie, P. (2007). Emotion. *Philosophy Compass*, *2*, 928-938.

- Goldie, P. (2009). Getting feelings into emotional experience in the right way. *Emotion Review*, 1, 232-239.
- Goldstein, I. (2002). Are emotions feelings? A further look at hedonic theories of emotions. *Consciousness & Emotion*, 3, 21-33.
- Goldthorpe, J. H. (1998). Rational action theory for sociology. *British Journal of Sociology*, 49, 167-192.
- Gordon, R. M. (1978). Emotion labeling and cognition. *Journal for the Theory of Social Behaviour*, 8, 125-135.
- Green, O. H. (1992). *The emotions: A philosophical theory*. Dordrecht: Kluwer.
- Griffith, P. E. (1997). *What emotions really are*. Chicago: University of Chicago Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Helm, B. (2001). *Emotional reason: Deliberation, motivation, and the nature of value*. Cambridge: Cambridge University Press.
- Hempel, C. G. (1965). *Aspects of scientific explanation*. New York: Free Press.
- James, W. (1884). What is an emotion? *Mind*, 9, 188-205.
- James, W. (1950). *Principles of Psychology* (Vol. 1 & 2). New York: Dover. (Original work published 1890).
- Köhler, W. (1947). *Gestalt Psychology* (2nd ed.). New York: Liveright.
- Külpe, O. (1893). *Grundriss der Psychologie. Auf experimenteller Grundlage dargestellt* [Outlines of psychology]. Leipzig: Wilhelm Engelmann.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50, 249-258.
- Lipton, P. (2004). *Inference to the best explanation*. London: Routledge.
- Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.
- Mackie, J. (1974). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.
- Madl, T., Baars, B. J., Franklin, S. (2011). The timing of the cognitive cycle. *PLoS ONE*, 6: e14803. doi:10.1371/journal.pone.0014803
- Mandler, G. (1984). *Mind and body*. New York: Norton.
- Marks, J. (1982). A theory of emotion. *Philosophical Studies*, 42, 227-242.
- Marsella, S., Gratch, J., & Petta, P. (2010). Computational models of emotion. In K. R. Scherer, T. Bänziger, & E. Roesch (Eds.) *Blueprint for affective computing: A sourcebook* (pp. 21-46). Oxford: Oxford University Press.
- Meinong, A. (1894). *Psychologisch-ethische Untersuchungen zur Wert-Theorie* [Psychological-ethical investigations in value theory]. Graz, Austria: Leuschner & Lubensky. Reprinted in R. Haller & R. Kindinger (Eds.). (1968). *Alexius Meinong Gesamtausgabe* [Alexius Meinong's complete works] (Vol. III, pp. 3-244). Graz, Austria: Akademische Druck- und Verlagsanstalt.
- Meinong, A. (1910). *Über Annahmen* [On assumptions] (2nd ed.). Leipzig, Austria: Barth. Reprinted in R. Haller & R. Kindinger (Eds.). (1977). *Alexius Meinong Gesamtausgabe* [Alexius Meinong's complete works] (Vol. IV). Graz, Austria: Akademische Druck- und Verlagsanstalt.
- Meinong, A. (1917). Über emotionale Präsentation [On emotional presentation]. In *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien, Philosophisch-Historische Klasse* (Vol. 183, pp. 1-181). Wien: Hölder. Wiederabdruck in: R. Haller & R. Kindinger (Hrsg.) (1968), *Alexius Meinong Gesamtausgabe* [Alexius Meinong's complete works] (Vol III) (pp. 285-465). Graz: Akademische Druck- und Verlagsanstalt.
- Mele, A. R. (1992). *Springs of action: Understanding intentional behavior*. NY: Oxford University Press.
- Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, 126, 910-924.
- Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, 8, 423-429.
- Meyer, W.-U., Reisenzein, R. and Schützwohl, A. (2001). *Einführung in die Emotionspsychologie, Band I* [Introduction to the psychology of emotions, Vol I]. Bern: Huber. (2nd edition)
- Miceli, M., & Castelfranchi, C. (1997). Basic principles of psychic suffering: A preliminary account. *Theory & Psychology*, 7, 769-798.
- Miceli, M., & Castelfranchi, C. (2002). The mind and the future: The (negative) power of expectations. *Theory & Psychology*, 12, 335-366.
- Miceli, M., & Castelfranchi, C. (2007). The envious mind. *Cognition and Emotion*, 21, 449-479.
- Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, 1, 29-50.

- Ortony, A., Clore, G. L. & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
- Pitcher, G. (1965). Emotion. *Mind*, 74, 326-346.
- Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.
- Pugmire, D. (1998). *Rediscovering emotion*. Edinburgh: Edinburgh University Press.
- Reisenzein, R. (1994a). Kausalattribution und Emotion [Causal attribution and emotion]. In F. Försterling & J. Stiensmeier-Pelster (Eds.), *Attributionstheorie: Grundlagen und Anwendungen* (pp. 123-161). Göttingen: Hogrefe.
- Reisenzein, R. (1994b). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, 67, 525-539.
- Reisenzein, R. (2000). Einschätzungstheoretische Ansätze [Appraisal theories]. In J. H. Otto, H. A. Euler, & H. Mandl (Eds.), *Emotionspsychologie: Ein Handbuch* (pp. 117-150). Beltz: Psychologie Verlags Union.
- Reisenzein, R. (2001a). Appraisal processes conceptualized from a schema-theoretic perspective: Contributions to a process analysis of emotions. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 187-201). New York: Oxford University Press.
- Reisenzein, R. (2001b). Die Allgemeine Hedonistische Motivationstheorie der Sozialpsychologie [The General Hedonistic Motivation Theory of Social Psychology]. In R. K. Silbereisen & M. Reitzle (Eds.), *Bericht über den 42. Kongress der DGPs in Jena 2000* (pp. 649-661) Lengerich: Pabst Science Publishers.
- Reisenzein, R. (2006). Arnold's theory of emotion in historical perspective. *Cognition and Emotion*, 20, 920-951.
- Reisenzein, R. (2007). What is a definition of emotion? And are emotions mental-behavioral processes? *Social Science Information*, 46, 424-428.
- Reisenzein, R. (2009a). Emotions as metarepresentational states of mind: Naturalizing the belief-desire theory of emotion. *Cognitive Systems Research*, 10, 6-20.
- Reisenzein, R. (2009b). Emotional experience in the computational belief-desire theory of emotion. *Emotion Review*, 1, 214-222.
- Reisenzein, R. (2010). Moralische Gefühle aus der Sicht der kognitiv-motivationalen Theorie der Emotion [Moral emotions from the perspective of the cognitive-motivational theory of emotion]. In M. Iorio & R. Reisenzein (Hg.), *Regel, Norm, Gesetz. Eine interdisziplinäre Bestandsaufnahme* [Rule, norm, law: An interdisciplinary survey] (pp. 257-283). Frankfurt am Main: Peter Lang Verlag.
- Reisenzein, R. (2012). Extending the belief-desire theory of emotions to fantasy emotions. *Proceedings of the ICCM 2012*, 313-314.
- Reisenzein, R., & Bördgen, S., Holtbernd, T., & Matz, D. (2006). Evidence for strong dissociation between emotion and facial displays: The case of surprise. *Journal of Personality and Social Psychology*, 91, 295-315.
- Reisenzein, R., & Döring, S. (2009). Ten perspectives on emotional experience: Introduction to the special issue. *Emotion Review*, 1, 195-205.
- Reisenzein, R., & Junge, M. (2006). *The intensity of disappointment and relief as a function of belief and desire strength*. Talk presented at the 45th Congress of the German Association of Psychology, Nuremberg, 2006.
- Reisenzein, R., & Junge, M. (2012). Language and emotion from the perspective of the computational belief-desire theory of emotion. In P. A. Wilson (Ed.) *Dynamicity in emotion concepts (Lodz Studies in Language*, 27, pp. 37-59). Frankfurt am Main: Peter Lang.
- Reisenzein, R., Meyer, W.-U., & Schützwohl, A. (1995). James and the physical basis of emotion: A comment on Ellsworth. *Psychological Review*, 102, 757-761.
- Reisenzein, R., Meyer, W.-U., & Schützwohl, A. (2003). *Einführung in die Emotionspsychologie, Band III: Kognitive Emotionstheorien. [Introduction to the psychology of emotions, Vol 3: Cognitive emotion theories]*. Bern: Huber.
- Reisenzein, R., & Schönplflug, W. (1992). Stumpf's cognitive-evaluative theory of emotion. *American Psychologist*, 47, 34-45.
- Roberts, R. C. (2003). *Emotions: An essay in aid of moral psychology*. Cambridge: Cambridge University Press.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145-172.

- Salmela, M. (2002). The problem of affectivity in cognitive theories of emotion. *Consciousness & Emotion*, 3, 159-182.
- Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 49-80). New York: Academic Press.
- Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, Methods, Research* (pp. 92-129). Oxford: Oxford University Press.
- Scherer, K. R. (2005). What are emotions? And how can they be measured?. *Social Science Information*, 44, 693-727.
- Siemer, M. (2008). Beyond prototypes and classical definitions: Evidence for a theory-based representation of emotion concepts. *Cognition and Emotion*, 22, 620-632.
- Smith, M. (1987). The Humean theory of motivation. *Mind*, 96, 36-61.
- Solomon, R. C. (1976). *The passions*. Garden City, NY: Anchor Press/Doubleday.
- Stumpf, C. (1899). Über den Begriff der Gemüthsbewegung [On the concept of emotion]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 21, 47-99.
- Stocker, M. (1983). Psychic feelings: Their importance and irreducibility. *Australasian Journal of Philosophy*, 61, 5-26.
- Whiting, D. (2011). The feeling theory of emotion and the object-directed emotions. *European Journal of Philosophy*, 1, 281-303.
- Wooldridge, M. J. (2002). *Reasoning about rational agents*. Cambridge, MA: MIT Press.
- Wundt, W. (1896). *Grundriss der Psychologie* [Outlines of Psychology]. Leipzig: Engelmann.