Emotion, Agency and Time: An Ontological Approach

Ombrato, Michele Davide

EMOTION, AGENCY AND TIME

an Ontological Approach

Michele Davide Ombrato

A dissertation submitted to the Department of Philosophy, University of Geneva in fulfilment of the requirements for the degree of Doctor of Philosophy in Philosophy.

November 2022

THÈSE DE DOCTORAT ÈS LETTRES

UNIVERSITÉ DE GENÈVE

Supervisor

Professor Julien Deonna (University of Geneva)

Co-supervisor

Professor Matthew Soteriou (King's College, London)

Jury members

Professor Michael Martin (University of Oxford, Corpus Christi),

Professor Lucy O'Brien (UCL, London),

Professor Naomi Eilan (University of Warwick)

President of the jury

Professor Fabrice Teroni (University of Geneva)

Contents

Acknowledgements

In the process of writing this thesis I have benefited immensely from the input and support of my two excellent supervisors: Julien Deonna and Matthew Soteriou. I want to thank them for their investment in my project, for their kind encouragement, and for the time and care they have dedicated to supervising my research over the years. I owe to them and their work many essential insights, the discovery (and the rediscovery) of my passion for philosophy, and much of my long-lasting interest in the topic of this thesis. I am also deeply grateful to Fabrice Teroni for advice, precious insights, and many interesting discussions, and for his generosity in providing me with comments on earlier drafts of several portions of this thesis. His input and his kind encouragement have been crucial for the completion of this work. Matthew Soteriou was my host during my stay in the UK at Warwick University and King's College London before he became officially my co-supervisor. This stay allowed me to take decisive steps towards delineating the project for this thesis. I am deeply grateful for the kind, thoughtful, unwavering encouragement I received from him at the time and for the very enriching opportunity to pursue my project at Warwick and KCL.

Much of the research for this thesis was conducted in the extremely stimulating environment of the Swiss Centre for Affective Sciences at the University of Geneva, and especially within Thumos, the research group on emotions co-directed by Julien Deonna and Fabrice Teroni. It has been a great pleasure and immense privilege to be able to work and discuss philosophy with Julien Deonna and Fabrice Teroni and with my doctoral and postdoctoral colleagues: Angela Abatista, Constant Bonard, Joffrey Fuhrer, Jacques-Henri Vollet, Steve Humbert-Droz, Roberto Keller, Arturs Logins, Edgar Phillips, Antoine Rebourg, Laura Silva, Magalie Schor and Juliette Vazard. I'm especially grateful to Edgar Phillips for his friendship and the

many enriching philosophical discussions. I am also grateful to Daniel Vanello. It was a great fortune to be able to share with him my interest in affect and emotion during my stay in the UK and I'm indebted to him for helpful discussions over the years.

I would like to thank the University of Geneva, the Ernst & Lucie Schmidheiny Foundation and the Société Académique de Genève for awarding me with grants which contributed to support my research in times of need allowing me to complete this thesis

Finally, I want to thank my parents Santa and Carmine, my little sister Federica, and my old friends from my years in Parma and Torino. Their emotional insight, their affection and their warm encouragement were truly invaluable throughout the process leading to the completion of this thesis.

Introduction

In this thesis, I provide an analysis of the causal structure of occurrent emotions understood as complex, episodic, dynamic psychological continuities over time. The broad aim of the analysis is to uncover and theorise the relations between the *temporality* of emotion episodes–i.e., their having temporal extension, boundaries, trajectories etc.–and their '*agentiality*'–i.e., the aspect under which they are ways or products of ways of being active.[1]

My main proposal will be that emotion episodes, unlike other kinds of affective reaction–such as, e.g., pain, sensory pleasure, physical craving etc.–are not wholly passive phenomena. Although they are elicited (or induced or triggered) in us–i.e., they are *initiated* from without–what causes them to persist and unfold as they do is an interplay between passive factors and agential factors, and more precisely, an interplay between 'affective changes'[2] and various agential cognitive processes–e.g., perceptual attending, imagining, ruminating, anticipating etc. These various processes constitute what I will call our 'sustained attentional engagement' with emotional situations, an *affective* species of *interest*-driven attending.

---

1 This formulation paraphrases Gordon (1986) on the traditional classification of "emotions as 'passions,' or ways (or products of ways) of being acted upon." (372). Cf. Peters & Mace (1961-1962).

2 The term 'affect' is used in the emotion literature in two main ways: (*i*) to refer to an overarching psychological category including phenomena such as, e.g., pain, sensory pleasure, thirst, hunger, fatigue, mood, and emotion; (*ii*) to refer to a particular component of such phenomena, namely the one which warrants their classification as *affective* phenomena or forms of affect. In this thesis, I will employ the term 'affect' in the second manner. The definition of 'affect' is object of controversy. I'll be assuming that 'hedonic valence', i.e., the phenomenal property of feeling 'bad' or 'good', is its essential characteristic.

To illustrate the proposal, on the general picture of emotion that will result from my analysis, passive and agential factors not only coexist, but produce one another. When we react with affect to given emotional situations, an affective interest is aroused and we actively engage with such situations as *objects* of our current interest and attention; we thereby sustain the continued elicitation of affective changes we undergo, and, hence, the renewal of the interest that affect arouses. It is this sort of cycle that underlies the *persistence* of emotion, and the many ways in which emotion episodes may *vary* as they unfold. The analysis of the structure of emotion episodes I will put forward has implications concerning the ontology of emotion, the nature of emotional cognition, and the nature of the relation that we bear to *value* while undergoing emotion episodes. In what follows, I will locate my thesis in relation to current theorising about emotion before I move onto outlining its structure. This will allow me to highlight the pertinence of my analysis to current philosophical debates as an attempt to remedy a neglect: the pervasive tendency to approach occurrent emotions while abstracting from their complex, episodic, dynamic character.

That occurrent emotions are complex, episodic, and dynamic is a largely agreed upon claim.[3] Emotions are complex in their being *syndromes* of changes in various aspects of mind–i.e., affect, or hedonic feelings, somatic feelings, attention, cognition and motivation–they are episodic in their having temporal boundaries or *onsets* and *offsets*, and they are, at least *typically*, dynamic in that they rarely–if ever–persist without variation. To get some grasp on the extent of what I mean here by 'variation', consider that, arguably, all emotion episodes undergo changes in intensity which seem to outline stages (or phases) of natural development–i.e., rise, peak(s), plateau(s), subsidence–while some emotion episodes undergo what I will call

[3] Cf. Mulligan & Scherer (2012, 346-347). This may seem to echo Goldie (2000): "[a]n emotion […] is typically *complex*, *episodic*, *dynamic*, and *structured*." (12). By 'an emotion', however, Goldie means what most would call an 'emotional disposition'–e.g., 'John's *love* for Jane' or 'Jane's *pride* for her child' understood as enduring attitudes. His use of the terms at issue differs from mine accordingly. He takes in fact such emotional dispositions to be complex, episodic, and dynamic in that they manifest themselves in disparate psychological episodes–including emotion episodes, but also moods, thoughts, perceptions etc.–that wax and wane over time (op. cit. 12-16). By contrast, I am interested in the complex, episodic, dynamic character of the emotion episodes themselves.

attitudinal transitions, or changes in attitude, such as, for example, shifts from 'shame' to 'regret' and 'dread'. To clarify, 'changes in attitude' are not mere 'sequences of emotions', such as they may occur while we react to currently unfolding situations–e.g., 'apprehension' or 'watchfulness' that turn into 'fear' or 'terror' as the threat materialises, or 'fear' or 'terror' that give way to 'relief' as the threat dissolves. Changes in attitude happen in reaction to unchanging emotional situations. They are due to the fact that disparate enduring concerns (Frijda, 2007)[4] are often at stake in any given situation, and therefore the emoting subject's ongoing *appraisal* of the situation's significance may over time *crystallize* in disparate attitudes, or so I will argue. Consider, to illustrate, how 'shame', 'regret' and 'dread' may succeed one another as one learns how badly one acted at a party after too many drinks: we can imagine the subject's attitude shifting from 'being *ashamed* of herself' on account of her actions to '*regretting* her drinking so much' to '*dreading* the prospect of facing certain people in the future.' In such cases, I suggest, it is natural to think of the emotional transitions we would describe by referring to different emotion attitudes as transitions undergone by single, *unitary* episodes of emotion as opposed to sequences of emotions.

Now, the idea that emotions are complex dynamic episodes, in the sense above clarified, arguably belongs to what we may call our naïve or *pre-theoretic* ontology of emotion. Indeed, the complex, episodic, dynamic character of emotion is commonly mentioned in philosophy and psychology amongst the *platitudes* that any plausible account of emotion is expected to accommodate.[5] Emphasis on these very basic ontological properties, however, has so far remained a bit *idle* in theorising about emotions, especially so in philosophical discussions.[6]

[4] Cf. Roberts (2005) for extended discussion of the relevant concept of 'concern'.
[5] See §2.2. Cf. e.g., Deonna & Teroni (2012, 8); Mulligan & Scherer (2012); Robinson (2005; 2018), Tappolet (2016, 3). Kuppens (2015) for the relevant references in affective sciences.
[6] In affective sciences, concerns about the lack of inquiry into issues connected to complexity and temporality of emotion have been voiced since the early days of 'appraisal theories'–cf. Scherer (1987), Frijda e*t. al.* (1991), Lewis (1995), Davidson (1998). They have become increasingly influential in theorising and empirical practice over the last two decades–cf., e.g., Scherer (2009) on the wide consensus in affective sciences on the superiority of *componential* approaches to emotions, and the growing scientific literature on the temporality of emotions, i.e., so-called 'affect dynamics.' See Kuppens (2015) and the other articles published in the Emotion Review, Vol. 7, No. 4, Special Section: *Affect Dynamics*, for an overview of the literature and further references.

First of all, notice that while the issue of whether emotions are best understood in terms of 'somatic feelings', 'judgements', 'perceptions', 'evaluative construals', 'transient attitudes' etc., usually takes centre stage,[7] the most basic or *fundamental* questions of ontology, questions which would naturally arise from keeping in sight the complexity and the temporality of these phenomena–"how do their aspects interact?", "What marks their temporal boundaries?," "What changes can they undergo as they persist?"–are still rather neglected.[8] One evident problem with the lack of inquiry into such questions is that while the idea that emotions are complex and dynamic episodes over time is commonly mentioned as one of the platitudes about them, there is deep (and often implicit) disagreement on how to understand these basic ontological properties. On closer inspection, emotion theories appear to diverge on just how complex emotions are–i.e., on which of the changes that take place during emotion episodes are *proper* parts of emotion, and which ones, instead, are better viewed as causes, effects, manifestations etc.[9]–on where to draw their exact temporal boundaries, and, relatedly, on just how much variation they can survive without *becoming* different emotions–or ceasing to exist as such while leaving behind residual arousal, somatic feelings or moods. As a result of such unarticulated disagreements, many emotion theorists identify what I will call their units of analysis, or 'emotional particulars', with episodes of emotion that overlap only partially–both synchronically and diachronically. Furthermore, while the basic questions of ontology of emotion are relatively marginal, the most central questions in current debates, namely (*i*) "what is the nature or essence of the emotions?", (*ii*) "What causes emotions?", and (*iii*) "How do emotions relate us to value?", are typically framed in ways which abstract from both their complexity and their temporality.[10] To wit, (*i*) the central aim of the vast majority

[7] Deonna & Teroni (2012) reconstruct in much detail this central philosophical debate. Cf. Tappolet (2016).
[8] Cf. Naar & Teroni, eds. (2018, Introduction).
[9] Cf. Soteriou (2018, 71).
[10] The most notable exceptions are Goldie (2000; 2012), Robinson (2005, 2018) and Wollheim (1999). See also Soteriou (2018) and Na'Aman (2019).

of accounts of emotion is to identify and theorise the essential component(s) of emotions[11]– as opposed to identifying and theorising the distinctively *emotional* interactions that take place amongst the changes which constitute them;[12] (*ii*) questions of causation about emotions usually regard their punctual initiatory elicitation as opposed to continued causation,[13] and (*iii*) the cognitive-evaluative relations that emoting subjects are taken to bear to the objects or contents of their experience are usually theorised in essentially *synchronic* or static terms– i.e., in terms of an *apprehension* of evaluative properties as opposed to an ongoing appraisal of emotional situations. This should suffice to motivate an attempt to diagnose the rationale for this sort of approach to emotions and remedy the neglect to which it tends to lead.

One key aim of this thesis is to propose that questions of ontology concerning specifically the complexity and temporality of emotion episodes ought to be prioritised so as not to risk individuating the basic units of analysis of emotion on grounds of unwarranted assumptions. One other aim is to show that addressing the central questions in current debates on emotion while we keep in clear sight the complexity and temporality of the phenomena is instrumental to uncovering a crucial aspect of our occurrent emotional life: its agentiality. As we will see, in fact, recognising and articulating the involvement of mental agency in emotion is essential to correctly theorise the way in which the various emotional components mutually interact, continued emotion causation, and the diachronic dimension of our being related to values while undergoing emotional experience. In what follows, I will outline the structure of this thesis and sketch the central arguments to be developed during its course.

The thesis consists in two parts. The aim of Part I is twofold: first, to elucidate the concept of 'sustained emotional engagement', and second, to make room for it by undermining the

[11] The influence of this sort approach is reflected in the centrality of the debate about whether emotions are *essentially* 'feelings', 'judgements', 'perceptions' and so on. See, e.g., Prinz (2004, 4) and Tappolet (2016, 8-15) for explicit endorsement of this approach. Cf., Goldie (2012, 64) and Soteriou (2018) for critical discussions.
[12] This sort of view of the nature of emotionality is often implied in the literature in affective sciences. Cf., e.g., Scherer's (2009) and other appraisal theorists' emphasis on *integration* of *synchronised components* as the defining feature of proper emotion episodes.
[13] Cf., Moors's (2009) reconstruction of current debates the '*elicitation* problem'.

motivations behind approaches which abstract from complexity or temporality–and, hence, obscure the agential aspects of emotion episodes. In chapter 1, I elucidate the concept of 'sustained emotional engagement' by elaborating on remarks put forward by Frijda as regards the relations amongst emotion, interest and (mental) agency. Frijda suggests that we can understand the various specific emotions as "variants of interest" (op. cit., 26), for all episodes of emotion, arguably, involve the subject's engagement with emotional situations as objects of interest and attention. He also suggests that 'engagement' is the most basic form of action readiness (Frijda 1986)–i.e., a form of action readiness underlying the *specific* action tendencies he associates to the various emotions.[14] I elaborate on these suggestions and argue for the initial plausibility and potential significance of the view that emotion episodes are *both* episodes of affective reaction and episodes of sustained emotional engagement. In chapters 2 and 3, I articulate my diagnosis of the pervasive tendency to approach occurrent emotions abstracting from complexity and temporality and try to undermine its rationale. In doing so, I build up to my analysis of the causal structure of emotion episodes. The proposed diagnosis traces back the tendency in question to two pivotal assumptions: (*a*) the assumption that the basic units of analysis for the study of emotion are stretches of emotional experience during the course of which one single attitude happens to be retained–i.e., they are episodes *of*, say, 'fear', or 'anger', or 'shame'–and (*b*) the assumption that occurrent emotions, as kinds of affective reaction, are *wholly* passive phenomena–viz., that they are through and through caused from *without*. In chapter 2, I raise the issue of the basic units of analysis of emotion and show that different accounts often identify such units with stretches of emotion that overlap only partially. I then deny (a) by showing that the concepts of 'emotion episode' and 'emotion attitude' are not necessarily coextensive–viz., that emotion episodes may *span* changes in attitude. I do so by arguing that we can individuate emotion episodes *independently* of attitude-*retention* and by offering a criterion of individuation for emotion episodes that may

[14] Frijda (1986). See Deonna & Teroni (2012, ch 7) for a philosophical development of this idea.

apply to both 'mono-attitude' and 'multi-attitude' emotion episodes. The key idea will be that we can individuate emotion episodes as transient affective changes demarcated by the way in which we deploy our attention while undergoing them. This dual criterion yields clear distinctions between emotion episodes and 'mere' emotional affect or mood. Furthermore, it provides us with a fresh point of departure allowing us to address questions of ontology of emotion while free from unwarranted assumptions concerning the basic units of analysis for the study of this phenomenon. In chapter 3, I raise the first one of such questions, namely, the question of emotional persistence. By addressing it, I build up to my analysis of the causal structure of emotion episodes as episodes over time. I start by examining the causal concept of 'reaction' with the aim to clarify what exactly classifying emotion episodes as kinds of affective reaction implies as regards their causal structure. The main upshot of the discussion is the rejection of the view that an affective reaction is as such a *wholly* passive phenomenon. There is in fact no conflict between the concept of affective 'reaction' and agential causation. My argument for this claim turns on the idea that the sort of *passivity* implied by 'reaction' pertains to the *punctual* initiatory elicitation of the phenomena, as distinct from how they are caused to persist or unfold over time. I then start to build up to my analysis of the structure of emotion episodes by identifying the main structural difference between emotion episodes and other kinds of affective reaction–especially pain-episodes. I start with an analysis of the role that attention plays across different affective reactions, including emotion episodes, and I argue that this role always remains functionally analogous: attention serves in all the cases the essential function of preparing and motivating action aimed at coping, or dealing with the affect-eliciting situations. Roughly, in all affective reactions, the attention is *continually* drawn to the reaction's elicitors; as a result, we feel *as if* we may not be able to fully attentively engage in any other (mental) activity requiring attentional resources until we have dealt with the affect-eliciting situation. I then show that the role of attention in emotion episodes is different, however, in two essential respects:

firstly, the *locus* of control of attention-orienting in other kinds of affective reactions is *exogenous*, whereas in emotion episodes it is *endogenous*–viz., attention is *interest*-driven. Secondly, our ongoing attentional engagement is not required for the persistence of our pain, thirst, hunger etc., whereas sustained emotional engagement is essentially required for the persistence of emotion episodes. Both structural disanalogies, I argue, are due to the fact that while other kinds of affective reaction are 'based' on bodily sensations, emotion episodes are mediated by cognition: sensations, in fact, may last whether or not attention is *drawn* to them, and, when attention is drawn to them, this is not *via* the arousal of an interest; by contrast, mental images and thought crucially depend on attention and, arguably, the ongoing *selective* occupation of attention by emotion-eliciting cognition does require the arousal of an interest. I conclude Part I of the thesis unpacking these ideas and discussing how the proposed ontological approach bears on the issue of the nature of the emotional subjective point of view.

During the course of Part II, I develop the analysis put forward in chapter 3 and detail the interplay of affect, affective interests and sustained attention. As I do so, I draw the main implications concerning the nature of sustained emotional engagement as mental agency governed by affective interests–viz., as endogenous yet affectively compelled. In chapter 4, I focus on (sustained) attending. The aim of the chapter is to specify the respect in which we exercise mental agency in episodes of sustained attending. I start by introducing a distinction between synchronic and diachronic attentional selectivity and showing that an account of the sort of mental agency we exercise in sustained attending must accommodate both kinds of selectivity and articulate their relations. This view connects with the idea that an account of attentional selectivity must capture the respect in which mental agency is exercised while leaving room for the passivity or *receptivity* of the mind as manifest, for example, in instances of 'noticing' and in the way in which specific thought contents *occur* to us while we are engaged in purposive cognition. In chapter 5, I develop the desired account of attentional

selectivity by elaborating on O'Shaughnessy's (2000) analysis of wakeful cognitive processes with the aim to extract precise criteria for sustained attending *vs* idle ongoing attending–e.g., 'daydreaming'. I then specify two sub-classes of sustained attending to which sustained emotional engagement arguably belongs. Sustained emotional engagement, I suggest, is exploratory or open-ended–i.e., it lacks completion conditions–and non-instrumental–it is sustained by a certain degree of *absorption* as opposed to an awareness of its instrumentality to achieving given purposes. Distinguishing sustained emotional engagement from unemotional exploratory absorptions leads me to identify different notions of 'interest', that which allows me to specify the *interest*-driven character of emotional attending. In the last chapter, I provide an account of the relation between (positive and negative) affective interest–and (positive and negative) 'value' and I contrast sustained emotional engagement as driven by affective interests with the sort of engagement with emotional situations that takes centre stage in Brady's (2013) approach to emotions–i.e., an emotionally motivated assessment of value occurring alongside our emotional reactions.

PART I

Chapter 1

EMOTION, INTEREST AND MENTAL AGENCY

1 Emotion episodes and emotion attitudes

In this first chapter, I specify the targets of my analysis–i.e., emotion episodes–by introducing a few key distinctions within the realm of emotion. Furthermore, I illustrate the key intuition behind my project–i.e., that emotion episodes are episodes of *sustained emotional engagement*–by drawing on various observations Frijda makes about emotions as episodes over time, interest, and (mental) agency. The aim of the chapter is to familiarise the reader with the way of thinking about emotion I will be articulating over Part I of this thesis more than persuade.

By 'emotion episodes' I mean those complex cognitive-*cum*-affective reactions we undergo–and *feel* as we do–that we usually identify or *describe* by means of terms like 'fear', 'guilt', 'envy', 'regret', 'anger', 'shame', 'sadness', 'jealousy', 'awe', 'pride', 'admiration', 'contentment' etc.–i.e., so-called 'specific' emotional categories. When using these terms to identify or describe the emotion episodes that we undergo, we identify or describe such episodes in terms of what I will call 'emotion attitudes', that is, by relating them to *evaluative* attitudes (that we take ourselves to have) held during their course. For example, in identifying how I feel whenever the neighbour's Rottweiler barks at me as 'fear', I identify a complex emotional reaction–i.e., a reaction which consists in various psychological changes across affect, attention, cognition etc.–by relating it to an attitude held towards the dog, namely an evaluation which represents

it as 'fearsome' (or 'dangerous', or 'threatening').[15] Analogously, in identifying how I feel when I learn about my bad drunken behaviour at last night's party as including *moments* of 'shame', 'regret', and 'dread', I identify my reaction by relating it to distinct attitudes held towards, say, my 'shameful' behaviour, my 'regretful' decision to drink so much, and the 'dreadful' prospect of facing certain people in the future. In this case, just like the 'fear' case, I identify a reaction to some *single* emotional (i.e., emotion-eliciting or 'concern-relevant') situation–i.e., the situation I'm in because I behaved how I did. It just so happens that this reaction, unlike the 'fear' reaction, goes through several attitudes directed towards distinct 'objects.' Nonetheless, my changing attitudes all stem from an ongoing process of appraisal directed towards the relevant eliciting situation,[16] an ongoing process which, over time, happens to *crystallise* in distinct emotion attitudes.[17]

I take this sort of characterisation of how emotion episodes and emotion attitudes are related to be neutral as regards the nature of emotion attitudes. It doesn't specify, for example, whether emotion attitudes are doxastic, conative, receptive, or *sui generis* evaluative attitudes.[18] It is not entirely neutral, though, as regards the nature of this relation; it excludes in fact the view that this relation is 'identity'–that is, that occurrent emotions *are*, at the same time,

---

[15] I remain mostly neutral as regards the metaphysics of evaluative properties, and more specifically, as regards the question of whether the relevant evaluative properties are response-dependent or response-independent (cf. e.g., Johnston 2001). In chapter 6, however, I will suggest that such values are more naturally conceived of as response-dependent properties if we are do justice to the specificity of emotional motivation from the point of view of the emoting subjects–e.g., sustained emotional engagement in fear is driven and governed by an interest in the situation as 'fearsome' *vs* 'dangerous.' Cf. Brady (2013) for an articulation of the contrasting view. See Tappolet (2016, ch.3) and Vanello (2018) for recent discussions and further references on these questions.

[16] My use of the phrase 'ongoing appraisal' in *non*-technical. I do not mean to refer to the sequence of appraisal 'checks' posited by several appraisal theories to explain the occurrence of emotional experience, or any other precise theory of appraisal processes understood as sub-personal processes (c.f Scherer & Ellsworth, eds. 2003). I will thus use the terms 'appraisal' and 'evaluation' quite interchangeably.

[17] Anticipating a bit, I understand the relations between the ongoing process of appraisal and emotion attitudes on the model of O'Shaughnessy's (2008, §17.3) account of the relations between 'trying to decide-whether' understood as "the active procedure of *trying to make up one's mind*" and its "inactive […] cognitive crystallisation" in an event of 'deciding-that'. Cf. Roessler (1999, 57). The analogy is, of course, partial, and solely structural, for there are many differences which stem from what I will call the open-endedness of appraisal and how mental agency is deployed in emotion.

[18] By characterising emotion attitudes as 'evaluative', I mean that they are attitudes connected to given 'thick' evaluations of their 'objects'). The notion of an '*evaluative* attitude' has become associated to Deonna & Teroni's (2012) account of emotion, or more precisely, to the idea that emotion attitudes are (thick) evaluative modes of presentation as opposed to representations of (thick) evaluative properties. I remain neutral on such an issue since nothing shall turn on it. On the notion of 'thick' *vs* 'thin' value see Williams (1985).

episodes and attitudes. That philosophers tend to think of occurrent emotions in terms of 'episodic' attitudes–as opposed to emotion episodes involving attitudes–is, as I will now try to show, quite clear.[19] To wit, although such a view is rarely explicitly endorsed, a number of theorists have felt the need to reject it. One of the most prominent advocates of introducing a clear distinction between emotion episodes and emotion attitudes is Robinson (2005). Robinson maintains that descriptions of emotion episodes in terms of emotional categories are (typically) retrospective folk-psychological assessments of their significance. To illustrate the relations between the emotion episodes we undergo and our self-ascriptions of emotion attitudes she provides the following example:

> If, as my boss leaves the room, I find myself trembling, with my fists clenching and unclenching, and my heart pounding, I may think to myself '*I must be* really resentful. *I must have* found it really unjust and offensive when he told me I didn't work fast enough.' In formulating these thoughts, […] I *hypothesize* that I must have seen my boss's behaviour as an affront, that I really didn't like it, and that it contravened an important goal of mine, such as maintaining my self-respect. In short, I subsume this event or series of events under the folk-psychological concept of 'resentment'. (op. cit. 81, italics added).

I do not agree with Robinson's suggestion that self-ascriptions of emotion attitudes are *just*, or even typically, retrospective assessments of emotion episodes–i.e., judgements "that [sum] up what I think I must have been attending to, what wants and interests I think I must have registered as being at stake, what thoughts I think I must have had about the unfolding situation" (ibid.). It is helpful, however, to start with such a view–which is quite extreme in its dismissal of emotion attitudes as constituents of the *experience* of the emoting subjects–to show that it is just not obvious that because we identify or describe our emotion episodes by

[19] See, e.g., Deonna & Teroni (2012). On the significance of the distinction between the view that emotions are attitudes *vs* involve attitudes, Milona & Naar (2021).

means of emotional categories occurrent emotions *are* 'episodic' attitudes. One of the reasons for which I reject this claim is that I wish to make room for 'multi-attitude' emotion episodes, i.e., episodes of emotion that *span* 'changes in attitude'–as opposed to being mere sequences of occurrent emotions. That such experiential continuities constitute emotional wholes or unities is in fact one of the central claims of this thesis. Moreover, besides the intuition one may or may not have that such phenomena are indeed unitary episodes, there are more general theoretical considerations that I think should incline us to avoid the identification of emotion episodes and emotion attitudes. As Deigh (1994) suggests, by identifying the emotion episodes we undergo with attitudes, we end up identifying *explananda* which harbour an internal tension between features the ascription of which is supported by their being episodes of emotion–that is the emotions' "turbulent" or experiential dimension–and features the ascription of which is supported by their being attitudes, that is, their intentional (evaluative) dimension (cf. op. cit., footnote 5). The idea here is that, on the one hand, "nothing in the general concept [of 'emotion'], as we ordinarily understand it, entails thought […]; nothing in our ordinary use of the words 'emotion', 'emotional', 'emote' or 'emotive' to describe a person's state of mind or behaviour implies that the person affirms or even just considers some propositions", (op. cit., 833) and, on the other hand, nothing in concepts such as 'fear', 'anger', 'shame' and so on seems to entail turbulence of the mind–as confirmed by the fact that we may also aptly use such terms to refer to 'cold' axiological judgements and standing emotional dispositions. This internal tension is normally resolved by identifying the emotion with one dimensions at the expense of the other, which, as Deonna & Teroni (2012), following Goldie (2000), rightly point out, ends up being theorised as a mere *add-on*: "[o]ne strategy involves complementing the *intellectual* judgment with a further layer of feelings […]. Another regards emotions as inchoate feelings given form by intellectual fiat" (*Preface* xi-xii).[20] Both sorts of add-on strategy have been criticised for being incapable of

[20] For further discussion of this approach, see Deonna & Teroni (2012, 56-57).

capturing the link between the *felt* and the intentional properties of emotions. This sort of criticism is one of the main motivations for those strategies which attempt instead to capture the intimacy of the link between phenomenal and intentional properties by proposing 'hybrid' notions such as, e.g., Goldie's (2000, ch.3) 'feelings-towards' and Deonna & Teroni's (2000) 'felt bodily attitudes.' I will not assess these various approaches. My own strategy will be to try and eliminate the internal tension that Deigh highlights by treating emotion episodes as complex episodes involving as one of their components the disparate emotion attitudes emoting subjects may come to hold, over time, *vis à vis* the relevant emotional situations. Later on, elaborating on Frijda's (2007) proposal that emotions are "variants of interest" (26), I will suggest that this evaluative-attitudinal component of emotion episodes is an *active* state of interest, namely a *mode of control* of attention: the emotion attitudes we come to hold during the course of emotion episodes account for the continuous selective occupation of attention by cognitive contents that are relevant to given 'thick' evaluations of emotional situations– e.g., as 'fearsome', 'offensive', 'shameful' etc. But for the time being, I just want to make it clear that I will not take emotion episodes and emotion attitudes to be identical: by employing emotional categories in ascriptions of emotion episodes we identify or describe such episodes in terms of the thick evaluative attitudes (we take to have been) held during their course.

So far, I have been concerned with distinguishing emotion episodes from emotion attitudes held while undergoing emotional experience. A further distinction is in order. It is often observed that emotional categories may also refer to standing, non-occurrent phenomena, that is, to the various enduring evaluative attitudes that typically underlie the occurrence of emotion episodes. Examples of this are one's standing admiration for the Italian Renaissance, fear of heights, envy of the rich, pride for one's children and so on. We usually invoke our enduring attitudes in the context of 'historical' or 'narrative' explanations of (patterns of)

emotion episodes.[21] One may, for example, mention one's standing admiration for the Italian Renaissance in an explanation of the intense admiration one *feels* whenever contemplating Michelangelo's '*Pietà*' in Saint Peter's Cathedral. In this sort of explanation, emotion episodes are primarily assessed as (appropriate) manifestations of one's enduring evaluative attitudes as opposed to (appropriate) reactions to worldly objects. Following Deonna & Teroni (2012), we may think of these phenomena as 'single-track' emotional dispositions. The label adverts to the fact that they dispose us to feel the corresponding occurrent emotions–for example, my fear (standing) of heights is my disposition to *feel* fear while I climb down a high cliff. Deonna & Teroni propose that single-track dispositions be kept distinct from 'multi-track' dispositions, e.g., 'love' and 'hate',[22] which instead dispose us to feel just about any kind of emotion depending on the precipitating circumstances–e.g., my love for my sister disposes me to feel, say, fear or anxiety if she faces some danger, pride or happiness if she has success in her life, and anger if she happens to be mistreated. We may now notice another problem with the idea that occurrent emotions are episodic attitudes: if we accept such view, we must then try and explain why terms such as 'fear', 'anger' or 'shame' happen to capture such ontologically disparate phenomena. This issue is rarely made explicit, but there seem to be two main solutions to be gathered in the literature: emotional categories signify primarily occurrent emotional phenomena and may derivatively be used to refer to their dispositional counterparts, or *vice versa*.[23] Although it is customary to simply clarify at the outset of inquiry whether one's *explananda* will be occurrents or dispositions, it is not clear that one may just gloss over the fact that these terms are aptly used to pick out both sorts of phenomena. Now, a natural explanation of this fact is that our emotional categories signify in both sorts of

[21] Cf. Wollheim's (1999) and Goldie's (2000; 2012) accounts of emotions, whose special focus on dispositions as opposed to episodes as the primary referents of emotional categories is supposed precisely to facilitate an historical or narrative approach to the phenomena. This sort of approach imports elements from the way emotions are theorised in Freudian psychoanalysis and other dynamic approaches to the mind. See Deigh (2001) on the legacy of Freud.

[22] These phenomena are also often called 'sentiments' or 'attachments.' Cf. Frijda (2007, ch.5); Shand (1926).

[23] The first option is by far the most widely endorsed. Prominent defenders of the second option are Wollheim (1999) and Goldie (2000).

emotional ascription the *same* sort of entity, that is, kinds of states, and more precisely, (doxastic, conative or *sui generis*) attitudes of subjects held towards given emotional situations. In the cases in which we are ascribing dispositions, emotional categories capture, so to say, the *entire* phenomenon. In the cases in which we are ascribing experiences, they fail to do so: they strictly signify an attitudinal-evaluative component of emotion episodes understood as syndromes which include ontologically disparate changes–i.e., changes both in state and experience. On this view, to pick out an episode of emotion as an episode *of* 'fear' or 'anger' is to identify it in reference to the emoting subject's continual retention of the attitude–i.e., as an episode for the duration of which a state of the kind 'fear' or 'anger' *continues* to obtain.[24] The view gains further plausibility if we accept what Deigh claims about the relations between the concept of 'emotion' and concepts such as 'fear', 'anger', 'shame' etc, namely that they do not have the same entailments; it is thus not obvious that we ought to think of the various emotions for which we have names as *species* of the *genus* 'emotion'.[25] Recognising that this is so may also explain why theorists who maintain that emotions are episodic attitudes face as one of their main challenges the challenge of reconciling the experiential and the attitudinal-evaluative dimension of their *explananda*.[26]

Having provisionally distinguished emotion episodes and emotion attitudes I may now turn to my positive characterisation of the former. As I have anticipated in the introduction, and as will see later in more detail, emotion episodes are 'complex' and 'dynamic' episodes, or, which the same, they are complex in two ways: synchronically (s-) and diachronically (d-); they are s-complex in that, at any given instant, they involve changes in various aspects of

[24] I owe this way of thinking of the relation between the episodic and the 'stative' aspect of emotion to Soteriou (2013, ch.2) and more specifically, to his concept of an *occurrent* state–i.e., a state which obtains for the duration of and in virtue of given experiential occurrences.
[25] Cf. Wollheim (1999, *Preface*, xii) for analogous observations.
[26] Cf. Deigh (1994): "[W]hat the relation is between the thought and the turbulence characteristic of emotional experiences and whether an emotion contains both thought and turbulent feeling as essential components or is to be identified with one of the two while the other is treated as an essential cause or a typical effect have become arguable questions, questions on which no consensus position has developed." (833, footnote 5).

mind, such as, e.g., *affect*, attention, cognition and motivation, and they are d-complex in that, as they persist or unfold over time, they may change in ways which mark temporal parts.

The idea that emotion episodes are s-complex is nowadays common currency,[27] but the idea that they are d-complex has not received nearly as much attention. As anticipated, I will understand the diachronic complexity of emotion episodes as follows. First, all such episodes are d-complex in that they undergo changes in *intensity* which outline typical stages (or phases) of development–i.e., rise, peak(s), plateau(s), subsidence. Secondly, some emotion episodes are multi-attitude episodes, i.e., they also involve changes in attitude or attitudinal transitions, e.g., a change from 'being afraid' to 'being angry' (Frijda *et al.* 1991; Frijda 1993; 2007, ch.7). Now, the fact that *all* episodes of emotion undergo changes in intensity suffices to support the claim that episodes of emotion are 'dynamic'. The 'simplest' and shortest episodes of emotion–e.g., very brief bouts of 'anger' elicited by someone stepping on our toes or 'fear' caused by a sudden crackle in the woods–and even what many would consider *sub-threshold* 'emotions'–e.g., 'startles' and what Frijda's calls 'small' emotions–possess intensity contours which outline stages of development (see Frijda 1994, 382; 2007, 185ff; cf. Robinson 1995). The number of further respects besides 'intensity' under which a discrete episode of emotion is, so to say, *allowed* to change as it persists crucially depends on the theorist's criteria of individuation for 'an' emotion. On some views, such as the one I'll put forward, an emotion may even span changes in attitude (Goldie 2000; 2012; Tappolet 2016, 3). On most views, however, changes in attitude are taken to mark temporal boundaries between separate occurrent emotions. I prefer the phrase 'emotion episodes' to the term '(occurrent) emotions' precisely because the latter, as I will try to show in chapter 2, is often reserved to one or more, but not all the various *synchronic* changes that constitute emotion episodes, and one or more, but not all the stages or phases that take place during their course. The debate to which

[27] Blackman (2013); D'Arms and Jacobson (2000, 2003, 2006); Goldie (2000, 2003, 2012); Greenspan (1988); Griffiths (1997); Marks (1982); Robinson (2005); Soteriou (2018); Stocker and Hegeman (1992).

this gives rise, namely the debate about which of the (synchronic and diachronic) parts of emotion episodes is to be identified with *the* emotion–as opposed to cause, effect or manifestation of the emotion–is one on which I remain neutral:[28] my analysis focuses on emotion episodes understood as emotional wholes or unities, and the aim of such an analysis is to identify and theorise the relations which obtain among their various components.[29]

As I have clarified in my introduction, my main aim in this thesis is to explore two distinct, but, arguably, closely related aspects of emotion episodes: their temporal dimension–that is, their having temporal extension, trajectories, and boundaries–and what I hope to show we may aptly call their 'agential' dimension, that is, their being ways, or products of ways, of being active. My main proposal will be that that these distinct aspects of emotion episodes are indeed closely related: emotion episodes are *partly* caused to persist and unfold over time as they do by various agential cognitive processes jointly constituting the emoting subjects' sustained emotional engagement. Understanding this phenomenon will take up most of this thesis as it involves an analysis of the ongoing interplay amongst affective changes, states of (positive and negative) affective interest and sustained attention. I provisionally characterise it as our being related to emotional situations as objects of interest and sustained attention–as opposed to elicitors of affect. In what follows, I will try to give some sense of the nature of sustained emotional engagement, and I will explain why I take it that it affords a vantage point of view on both the temporality and the agentiality of our occurrent emotional life.

2 Emotions as "variants of interest"

The idea that during emotions we are related to emotional situations not only *qua* elicitors, but also *qua* objects of interest and *interest*-driven attention is not new. It has been suggested

[28] Cf. Prinz, 2004 and Naar & Teroni, eds., (2018).
[29] The viability of this sort of approach is defended by Soteriou (2018, 74-75).

by various authors, namely Frijda (2007), Izard (1977), Johnston (2001) and Stocker (1987).[30] But Frijda's remarks on the relation between emotion episodes and engagement are the primary sources of the driving intuition of this thesis. As we will see, he explicitly connects, in fact, the idea that emoting subjects *engage* with emotional situations as objects of interest and attention with questions regarding the temporality and the agentiality of emotion. Before I turn to this connection, let me briefly outline Frijda's approach to emotions and the role interest and engagement play in such an approach.

Frijda (1986) has famously introduced in contemporary psychological and philosophical discussions an influential to approach to emotions built on the idea that such phenomena involve so-called *modes of action readiness*. A mode of action readiness is defined as a *tendency* to achieve a certain *relational* purpose, that is, establishing, maintaining, or interrupting a certain relationship with one's environment (op. cit., 71). The defining feature of such tendencies, as distinct from other *non*-emotional behavioural tendencies, is their *urgency*, or their enjoying "action control precedence" (op. cit. 78): urgency, or action control precedence is manifest in the emotions' tendency to interrupt, or pre-empt mental activity and purposive behaviour *irrelevant* the pursuit of the "relational purposes" that they, so to say, 'set for us', and, conversely, what Frijda calls the "indistractability" (op. cit. 460) of the emotional behavioural tendencies–i.e., their persistence in presence of potential distractors. We can say that for Frijda's emotions are mental and bodily preparations geared to action. As Scarantino (2005) puts it, on Frijda's approach: "emotions are *action control structures* which *prioritize* the pursuit of certain relational goals, prepare for their fulfilment both mentally and physically, and protect the execution of actions aimed at fulfilling them from possible interferences" (213, italics added). Frijda articulates this approach by arguing that different specific emotions involve different 'modes' of action readiness, that is, special action tendencies that are fit to

[30] Cf. also Brentano's (2009, 152-154) characterisation of emotions as instances of (positive or negative) interest. Ryle's (1954) account of *enjoyment* in terms of interest and Gallie's (1954) elaboration of it are also very pertinent. See Textor (2017, §3.3) for illuminating discussions of this literature. See ch.6 below.

answer the *demands* of their typical eliciting situations. To illustrate, episodes of 'fear' involve tendencies to act so as to deal or cope with dangers, for example, by achieving one's own inaccessibility, episodes of anger involve tendencies to act so as to deal or cope with offenses, for example, by holding one's ground, or retaliating, and so on. The idea that our specific emotions involve *distinctive* modes of action readiness is the most seminal element of Frijda's legacy; much theoretical effort has been put into its detailed elaboration.[31] Frijda's account, however, contains another important element that has not received nearly as much attention, namely the idea that each mode of action readiness is a distinctive *modulation* of a *shared* and more basic form of readiness, that lies beneath the specificity of the special action tendencies. Frijda (1986) calls this more basic form of readiness "readiness as such". In his first book, Frijda is not very forthcoming as regards the nature of 'readiness as such', but in his 2007 book he tries to clarify it in terms of the concepts of 'interest' and 'engagement': the modes of action readiness, he suggests, while differing from each other in their being tendencies to engage in given kinds of action, all presuppose a more basic form of readiness consisting in the emoting subjects' interest in, or engagement with, emotional situations: "the individual is engaged in what is going on and is interested by it. Emotions–all emotions–involve *engagement as the most general form of action readiness*. They are *variants of interest*. [Interest] may well be *the* basic emotion." (op. cit., 26, some italics added).

The appeal of this sort of concept in view of an analysis of emotion episodes such as the one I aim to put forward is quite clear: it is extremely congenial to the view that we may identify emotion episodes *independently* of their being episodes *of* given kinds of emotion in that it suggests that we may identify them as bounded episodes of sustained emotional engagement. Moreover, the idea that all emotions–viz., emotion attitudes–are "variants of interest" may

[31] In philosophy, the idea that specific emotions differ in their action tendencies has been fruitfully developed by Deonna & Teroni (2012) in an analysis of the evaluative epistemic dimension of emotions as felt attitudes. Scarantino's (2014) theory of emotions as action control systems also develops Frijda's idea focussing especially on the nature of emotional motivation.

be developed to try to elucidate the relation between emotion and mental agency, as I will presently illustrate.

First of all, notice that the idea that emotions involve engagement with emotional situations fits very nicely and naturally complements the view that emotions are action control systems: our being related to emotional situations as objects of current interest and *interest*-driven attention is arguably essential to an ongoing control of our actions; it is the mental aspect of our readiness to deal or cope with emotional situations, and more precisely, it affords the diachronic attentional selectivity which would seem to be required in order to avoid distraction and actually reserve control precedence to relevant courses of action over time.[32] The first and most obvious connection amongst emotion, (mental) agency and time inheres thus in the role of the emotions in ongoing action control. A less obvious, but more intimate connection–and the connection I'm mainly interested in exploring–is the one we may capture if we construe sustained emotional engagement as being itself an instance of mental agency–and more precisely, an instance of agential control.[33] Frijda did not pursue this idea; however, he did make relevant remarks I will now rehearse.

Frijda (2007, ch7) notices that emotion episodes identified as reactions to single *unchanging* emotional situations are often described by the subjects of self-report studies as involving sequences of various specific emotions. Interestingly, this seems to be taken by the subjects to be compatible with describing such episodes as wholes or unities falling under single emotion concepts. For example, episodes of the kind 'sadness' are described as including "despair, anger, and bewilderment, besides sorrow and distress" and episodes of 'jealousy'

[32] The features attributed to emotions as action control structures by Frijda (1986, 2007) and Scarantino (2005) under 'action control precedence' are features commonly attributed to cognitive control functions – see, e.g., Monsell & Driver, eds., (2000); Cooper (2010) and the other articles published in Topics in Cognitive Science Vol. 2, No. 4. That this is so since emotion and cognitive control functions are actually integrated neurologically is a more recent, non-mainstream development connecting research on emotion and attention to research on cognitive control functions (cf., e.g., Gray *et al.* 2002; Gray 2004 and Pessoa 2009; 2013).

[33] I call agential control the sort of agency we exercise over the way experiential continuities *unfold* over time as opposed to the one exercised, at given times, in *punctual* mental actions–e.g., judging, assenting, supposing etc.

as including "distress, anger and sheer anguish, besides […] pangs of jealousy." (op. cit. 181). Now, one way to make sense of such descriptions is to say that while a *single* episode of sustained emotional engagement is unfolding, the subject's attitude towards the relevant situation changes.[34] As regards the issue of what–if anything–underlies the subject's classification of such multi-attitude episodes as episodes of one kind of emotion, I suggest that it may warrant distinct explanations, case by case, and that there is no need to simplify matters by arguing for one general solution. To illustrate, in some cases we may be able to identify one single concern, such as, for instance, a single-track emotional disposition–e.g., one's standing 'jealousy' for one's lover–as lying behind the entire episode; in other cases, the global emotional categorisations the subjects offer might reflect what Robinson calls 'after-the-fact summary judgements' of emotional reactions. As Robinson (2005, 80-81) emphasises, in all cases the subjects' labeling is prone to errors. Let's consider an example Robinson provides.

> Suppose […] that my long-time husband were to abandon me for a younger woman. In such a situation my emotions are likely to be in turmoil: streams of grief, anger, shame, and despair intermingle and it would probably be hard for me to summarize my experience in a single word: the emotion process would be too complex and too ambiguous. After a while, when I catalogue the experience in recollection, I may say that I am *indignant* that my husband abandoned me for a younger woman, but in fact my behaviour and my physiological reactions reveal that my primary emotions are *shame* and *grief*. (op. cit. 81).

An emotional situation such as the one considered by Robinson is likely to be relevant to disparate concerns, and the fact that such concerns are aroused or *activated* in sequence, and,

[34] As I have clarified in §1.1, in such cases, the emotion attitudes are directed towards distinct 'objects.' However, in so far as the attitudes all stem from an ongoing process of appraisal directed towards one situation, for simplicity, I will speak of changing attitudes *towards* given emotional situations.

sometimes, simultaneously, plausibly explains the complexity our occurrent emotional life exhibits when we look at it through different 'time-frames'–e.g., from the order of seconds or minutes to the order of years.[35] Nonetheless, the situation, intuitively, has its own unity, and the emotion episodes one may undergo when thinking about it, I suggest, might each be viewed as unitary *explorations* of its relevance to one's concerns. Of course, much more needs to be done to defend such a view of emotional transitions as multi-attitude emotion episodes, but if a story along these lines were basically correct, the way in which ongoing episodes of engagement with emotional situations unfold over time would reveal that some, if not most of our occurrent emotional life, is made up of emotion episodes outstripping the boundaries of attitude-*retention*. Now, as we shift our focus onto these more 'articulate', multi-attitude episodes of emotion, a number of questions regarding the *temporality* of emotions come into view which remain somewhat concealed if we focus instead on single, specific, mono-attitude emotion episodes: in what does their identity over time consist, if not in their being emotion episodes during which an attitude is retained? What single causal process or constant manner of production underpins the way in which they *unfold* over time? What marks their temporal boundaries? As we will see, addressing such questions naturally leads us to look beyond the passive aspects of the emoting subjects' relation to emotional situations, aspects that are usually captured through the concept of emotion as 'reaction', and it lets into view agential aspects of this relation, that such concept–or better a certain understanding of it–somehow obscures. I take it that Frijda is suggesting something along these lines in the quote below.

> Emotions should *not* be primarily understood as *reactions*. They are best viewed as modulations of a prevailing background of continuous engagement with the environment. Fear *evoked* by a sudden *stimulus*

[35] As I will clarify in chapter 2, the relevant 'time-frame' in view of my purposes is somewhat intermediate. Unlike most chosen time-frames, it lets come into view multi-attitude emotion episodes, but, unlike, Goldie's, for example, it does not let come into view narratively structured sequences of 'emotions' ranging over days or years.

> is not the paradigm of emotion–not in humans and not in mammals that *explore* and run around. *Interest* and desire are better paradigms and so are lust and love. (Frijda 2007, 27, italics added).

In the same spirit, in an earlier paper, Frijda insists that episodes of "[g]uilt, anger, shame, and sadness *rarely fall upon an unprepared mind*. They build from previous doubts, irritations, ruminations, and apprehensions […]. [These various processes] are the *background* for what actually triggers the emotion." (1993, 382, italics added). The idea here is that the emoting subjects' contribution to the causation of emotions involves not only their standing concerns or sensitivities to given values, but also pre-existing backgrounds of mental activity. Because of such an ongoing mental activity, Fijda insists "emotions rarely have the one-shot, immediate and fast character of the paradigmatic case of being startled by a crackle in the solitary woods. And even there one has been walking around with all sorts of expectations against which the crackle is perceived." (ibid.). Now, the emoting subject's mental activity does not end with emotion *elicitation*; rather, it goes on as the emotion unfolds over time, and, crucially, it may evolve in ways which in some cases preserve while, in other cases, modulate our initial emotional reactions. Indeed, the way in which our mental activity evolves over time may also alter radically our initial reactions, resulting in articulate, multi-attitude emotion episodes.[36] Frijda insists that such instances of emotion are just as natural *explananda* of an account of emotion as the simpler, mono-attitude instances. To this I will add that multi-attitude emotion episodes turn out to be most *relevant* basic units of analysis if one has a special interest in the temporality and agentiality of our emotional life: they raise questions regarding the temporal character of emotion episodes which remain otherwise concealed,

[36] Cf. Frijda (1993): "[m]ost emotions have the form of protracted transactions (Lazarus 1991) or [complex] emotion episodes (Frijda, Mesquita, Sonnernans, & Van Goozen, 1991). And articulate emotional experience, as Ellsworth (1991) and Parkinson and Manstead (1992) have argued, is quite generally the outcome of successive steps of information pick-up that gradually build up that experience. Each step is determined by the outcomes of the preceding steps; elaborations are among those outcomes. [T]hese outcomes as well as previous doubts, ruminations, and relevant experiences coalesce, as it were, into a cognitive structure that serves as the input of the emotion-governing appraisal process at any given moment." (383).

and addressing such questions, as we will see, leads rather naturally to recognising the role of mental agency in emotion. Despite the interest of Frijda's remarks about emotion episodes and engagement, the implications, and possible theoretical advantages of their fully-fledged elaboration in an account of these phenomena are mostly unexplored. This thesis is partly an attempt to fill this gap. In the next section, I conclude this chapter by discussing how the recognition of the role of mental agency in emotion may bear on the nature of the link between emotion and evaluation.

3 Emotions and evaluation

Pursuing the claim that emotion episodes are episodes of sustained emotional engagement, besides offering a vantage point of view on the temporality and agentiality of the phenomena, will allow us to partially rethink the relation between emotions and the *evaluative* domain.[37] The central premise of the (nowadays orthodox) 'cognitive approach' to emotions, to which I subscribe, is that emotion episodes involve so-called 'appraisals' of the concern-relevance or significance of emotional situations. The exact nature of emotional appraisals is object of controversy. In philosophy, it has long been customary to think of them as *apprehensions* of ('thick') values or evaluative properties–e.g., 'fearsomeness', 'shamefulness', 'admirableness.' In psychology, 'appraisal process theories' propose an analysis of such value apprehensions into less conceptually sophisticated, 'molecular' appraisals, such as, for example, appraisals of (degrees of) 'novelty', 'pleasantness', 'goal-conduciveness', and 'coping potential'–cf., e.g., Ellsworth & Scherer (2003). What I wish to note here is that whatever the details of the view that emotions are or involve appraisals, if the relation that emoting subjects bear to value is an apprehension of evaluative properties–i.e., some kind of cognitive *achievement*–such a view suggests that this relation is a passive matter, and more specifically, a matter of receptivity.

[37] As introduction to the issue of the relation between emotion and evaluation, see Deonna & Teroni (2012).

Apprehending evaluative properties, an instance of knowledge acquisition,[38] does not itself involve any active contribution on the part of the emoting subjects; all that is required of them is that they harbour the various concerns which make up their evaluative sensitivity.[39] One's concern for one's physical and personal integrity, for example, is what plausibly constitutes one's sensitivity to 'fearsomeness', a sensitivity which manifests itself during episodes of 'fear', 'dread' or 'terror'. Now, while I take this picture to be correct, I also take it to be incomplete. Frijda's remarks are suggestive of the idea that our being related to evaluative properties during emotion is not merely a matter of *receptivity*; it also partly consists in ongoing *exploration* of the value (or relevance or significance) of the emotional situations. To substantiate this idea, we need to meet the challenge of reconciling the passive aspects and the agential aspects of emotions. I will briefly discuss what such challenge amounts to before I outline the solution that I will develop during the course of Part I of this thesis.

Although there are several commonly recognised connections between emotion and agency, especially in relation to conative features of emotion,[40] the idea that emotions are *themselves* agential is rather unorthodox.[41] The received wisdom is in fact that emotions are instances of passivity, 'passions' in the classical meaning, that is, "ways (or products of ways) of being

[38] Cf. Martin (2001, 265-266); Mulligan (2007, 213-214).

[39] This issue is orthogonal to the question of whether such value apprehensions occur prior to the emotions. On those accounts on which emotions are reactions to prior value apprehensions–e.g., to *being struck* by value–they are not *themselves* receptive aspects of the mind. Nonetheless, our being related to value during emotions remains a matter or receptivity–it still depends on apprehension. (Cf. Mulligan 2007; 2009; Müller 2017; 2022).

[40] Cf., e.g., Deonna & Teroni (2012); Frijda (1986); Scarantino (2005); Tappolet (2016).

[41] Müller (2021) has recently argued that emotions, as reactions having *motivating* reasons, are forms of *spontaneity*, and that they are in this respect to be contrasted with perceptions as forms of *receptivity*. Moreover, he explicitly says that "in thinking of emotions as felt for motivating reasons we conceive of them as a form of activity" (3). It is not clear, however, that his argument supports these claims. First of all, that spontaneity and receptivity should be thought of as standing in opposition to one another, as Müller assumes, is not obvious. McDowell (1996), for example, understands 'spontaneity' as a "label for the involvement of conceptual capacities", and cinsists that, although our conceptual capacities are capacities for "active" thought, "when we enjoy experience [they] are drawn on *in* receptivity, not exercised *on* some supposedly prior deliverances of receptivity. Indeed "it sounds off key in this connection to speak of *exercising* conceptual capacities at all. That would suit an activity whereas experience is passive" (10, cf. 9-13). In any case, as it becomes clear towards the end of the paper, Müller thinks of the respect under which he takes emotion to be a form of activity on the model of the evidence-coerced 'assent' involved in the formation of doxastic attitudes and this is not what I mean by agential aspects of emotion. Furthermore, Müller thinks of emotion as reactions to apprehensions of value by which they are preceded, so that the relation to value remains a matter of receptivity (see previous footnote).

acted upon" (Gordon 1986, 372), and not ways or products of ways of being active. Now, the common classification of occurrent emotions as kinds of affective reaction might seem to confirm this wisdom: reactions are 'elicited', 'induced', or 'triggered' *in* us by their specific elicitors, as opposed to being brought about at will. In this respect, emotions are just like sensory feelings–e.g., pain and sensorial pleasures–and perceptual experience. Furthermore, like these phenomena, they are often taken to play an epistemic function, that is, to give us some *access* to that with respect to which, while undergoing them, we are passive–namely, bodily states in the case of sensory feelings, various sensible qualities of worldly objects in perceptual experience, and the evaluative dimension of the world in the case of the emotions. That emotional epistemic functions necessitate the emoting subjects' passivity is insisted on by various authors.[42] Brady (2013), for instance, states that our passivity *qua* emoting subjects is strictly required if emotional appraisals are to be an "epistemically respectable affair" (44).

On the other hand, there are also important differences demarcating emotions from mere sensory feelings and perceptual experience; in particular, emotions are *responsive* to reasons, and not 'a-rational' phenomena.[43] This is shown by the fact that we commonly and intelligibly praise or blame people on account of how they (fail to) react emotionally to given situations, and that we speak of emotional reactions as the sort of state for being in which one may have one's own reasons–contrast, for example, "I have my own reasons for being angry" with "I have my own reasons for being warm (or being thirsty or seeing orange)." As it is the case with other clearly reason-responsive phenomena–e.g., doxastic attitudes–attributions of praiseworthiness and blameworthiness on account of emotional reactions seem to imply that we are to some extent responsible for how we react, and this, in turn, is often taken to imply that we must be able to exercise some kind of control over such reactions. This idea might gain further support from cases in which our emotional reactions seem to result from how,

[42] Cf. e.g., De Sousa (1987); Brady (2013); Prinz (2004); Roberts (2003); Tappolet (2016).
[43] For discussion of this distinction, see Scanlon (1998, ch.1).

given our aims, we *wish* things stood evaluatively as opposed to how they reasonably stand (cf. Griffiths 2008; Sartre 1939; Solomon 1976). The most often cited example of such cases is working oneself into a state of anger to enhance one's moral standing in a lovers' quarrel. The fact that we blame or praise each other on account of our emotions (or absence thereof) and that emotions might serve *strategies* aimed at (intrapersonal/interpersonal) manipulation are the intuitive sources of those theories in which the passive and epistemic-receptive character of emotions is outright rejected, and emotions are treated as 'actions' or 'choices,' as opposed to reactions (cf. Solomon 1976). It is widely accepted that such theories are not tenable as general theories of emotion (see, e.g., De Sousa 1987, 42; Roberts 2003, 31-36). Their initial appeal stems from focus on rather specific instances of emotion that cannot be considered paradigms,[44] and even such cases, in so far as they are to be cases of genuine emotion as opposed to pretend emotion, must involve some passivity. In order to see this, it should suffice to notice that the strategies by means of which one may work oneself into a fit of anger are not different from those by means of which one may induce a fit of anger in another person. In both cases, what one can produce is the presence to the mind of the targeted subject–i.e., oneself or the other person–of *triggering* thoughts, but in the end it is the trigger which, if apt, will cause the targeted subject to react emotionally.[45] Agential theories do, however, do justice to compelling intuitions that one may wish to preserve, intuitions that accounts of emotions as entirely passive aspects of the mind seem ill-suited to capture. The idea that an account of emotion must be able to fit both the passive aspects and the agential aspects of emotion has been taken seriously, among others, by De Sousa and Goldie. De Sousa (1987, 10) refers to it as the apparent "antinomy of activity and passivity" and Goldie (2012, 56) as the complex "pattern of activity and passivity, inner and outer" distinctive of emotional experience. In the same spirit, I suggest that one may accommodate

[44] Cf., Griffith (2008), who sets them apart as belonging to the special category of *strategic* emotions.
[45] On this, see Roberts' (ibid., 33-34) insightful discussion of the *got-up* emotions of Dostoevsky's underground man and the acting-method developed by Stanislavski. See also De Sousa (1987) on what he calls *bootstrapping*.

the coexistence of passive and agential aspects in emotion episodes. I will try to do so arguing that emotions are *both* episodes of affective reaction–i.e., episodes of *affection* by/reaction to emotional situations–and episodes of sustained emotional engagement. To let into view the passive aspects and the agential aspects of emotion, and their ongoing interplay over time we must consider emotion episodes in all their synchronic complexity and from a truly diachronic point of view, that is, as composite processes unfolding over time. In the two chapters that complete this first part of the thesis, I will reject two very common assumptions which represent obstacles to doing so: (*i*) the assumption that the basic units of analysis of emotion are to be identified in reference to single specific emotion attitudes and (*ii*) the assumption that emotion episodes, as kinds of affective reaction, are wholly passive.

Chapter 2

EMOTION EPISODES AS UNITS OF ANALYSIS OF EMOTION

Introduction

It is often observed that research on emotion in all disciplines is hampered by profound disagreement on how to *define* 'emotion."[46] One might think that this sort of observation shouldn't concern philosophers as much as psychologists and affective scientists. However, even though definitional disagreement is of course widespread and quite the norm in all areas of philosophy, in the case of the debate on defining 'emotion' it is so profound as to cast doubt on the unity of the topic. Wollheim (1999) expresses this sort of worry at the beginning of his book on the emotions.

> If we start on the most interesting, the most often cited, writers who are regarded as having written on the emotions we can come away with the idea that there are as many topics lurking under one word as there are writers addressing it. It is not simply that there is a lot of disagreement. I do not know that there is. *What is disturbing about the literature is the divergence in starting-point and the infrequency of any intersection.* (Wollheim, 1999, Introduction, xiii, italics added)

[46] Cf., Dixon (2012); Gendron (2010); Izard (2010); Kagan (2010); Russell (2012); Mulligan & Scherer (2012); Scarantino (2012).

Supposing that Wollheim is right on this, the situation would seem to require a special diagnosis. Let me elaborate a bit on Wollheim's observations. On one common conception of definitional disagreement, the latter requires a shared *definiendum*. In the case of the debate on how to define the concept of 'emotion', one may thus expect the different accounts, which greatly diverge as regards the precise nature of the phenomena–e.g., on whether emotions are somatic feelings, judgements, perceptions, motivations, *sui generis* attitudes etc.– to agree at least on how to preliminarily (provisionally) individuate their target phenomena. What I want to suggest is that this is precisely what Wollheim is doubting in making reference to the divergence in starting-point and infrequency of intersection in the emotion literature. In the same spirit, I will argue that much disagreement in the debate on defining 'emotion' is not definitional, in the sense above clarified, but, rather, ontological, i.e., pertaining, ultimately, to the way in which occurrent emotions are understood in connection to the most general or abstract categories we may apply to aspects of the mind, such as, for example, 'state,' 'event,' 'process,' 'episode,' 'occurrent,' 'disposition,' 'part,' 'whole,' 'boundary' etc.[47] Although emotion theorists rarely elucidate the way in which occurrent emotions are to be understood in connection to these basic ontological categories, it can be shown that distinct implicit commitments inform the concept of 'an' emotion that they employ, i.e., the way in which they *particularise* the stream of affect. The key idea I will articulate in this chapter is that our occurrent emotional life is in fact made up of experiential continuities that one may 'carve up', or delimit, in different manners, i.e., by means of distinct criteria of individuation. Since many accounts of emotion do employ different criteria, they cannot be taken to share the same *definiendum*. To avoid conflating the definitional and the ontological disagreement, I will call the various emotional occurrences one may individuate as one's target phenomena

[47] Cf. Mulligan & Scherer (2012, 347). In defence of ontology-based approaches to questions in the philosophy of mind see, e.g., Steward (1997; 2013); Soteriou (2013; 2018). On the pertinence to questions in the philosophy of mind of the category of 'process' in particular, see Stout (ed., 2018). On ontology of emotion in particular, see Naar & Teroni (eds., 2018).

'emotional particulars.' This theory-neutral terminology is useful in raising and addressing the 'individuation' question about emotion.

I will start to raise this question by illustrating the idea that we may carve up our occurrent emotional life in different manners, and more specifically, by showing that emotion theorists, by focusing on different features of emotion they take to be essential or defining features, end up individuating emotional particulars that *partially* overlap. I will then argue that since many of the features that different emotion theorists consider essential track different components of emotion episodes, starting the inquiry on emotion by offering an analysis of emotion episodes in their complexity is an attractive approach. I will then turn to the main aim of the chapter: to identify criteria of individuation for emotion episodes independent from attitude retention. Anticipating my conclusion, I will put forward a dual, conjunctive criterion: emotion episodes are temporally bounded alterations of our neutral or 'default' affective state as well as episodes of sustained attention. I will begin to build up to this conclusion by presenting the idea that emotion is *inherently*, or naturally episodic–it *comes* in separate episodes–and suggesting that this idea is part of what we may call our 'naïve' ontology of emotion. Elaborating on this idea, I will describe emotion episodes as bouts of affect with *natural* temporal boundaries and *natural* temporal trajectories. After this very abstract preliminary characterisation, I will further qualify emotion episodes by distinguishing them from other affective episodes, and more specifically, specific mood changes such as, e.g., bouts of 'elation,' 'anxiety,' 'irritability', 'depression' etc. To do so, Firstly, I will argue that emotion episodes involve specific mood changes as one of their components, and that, typically, they leave such changes behind in their aftermath. A corollary of this view is that emotion episodes differ from specific mood changes as wholes differ from their parts. Secondly, I will contrast emotion episodes and specific mood changes occurring in isolation. Drawing on remarks made by Pears (1962) on the relation between emotions and attention, I will suggest that emotion episodes require continual *cognizance* of their 'objects' as ensured

by sustained attention. I will finally present reasons in support of an account of emotion including more articulate, multi-attitude emotion episodes.

1 Delimiting emotional experience

The account of emotions developed by Peter Goldie (2000; 2012) is an ideal starting-point in view of raising the question of emotion-individuation. Although Goldie ultimately construes the emotions as kinds of 'state' understood as non-occurrent, dispositional phenomena, he also provides an account of the experiential continuities in which such states become manifest. I suggest that we may treat Goldie's account of these continuities as an account of his emotional particulars. What makes Goldie's emotional particulars so well suited to my current purpose is that they seem to be especially long-lasting compared to the ones featuring in most other accounts. We may thus view them as sort of maximally extended emotional particulars that most other accounts, in individuating their emotional particulars, further delimit.

Goldie's emotional particulars are experiences individuated by their *embeddedness* into a 'narrative structure'–roughly, a 'story-like' sequence of events, actions, feelings, and thoughts. Different emotional experiences are embedded in different *narratives* of parts of our lives, and it is in the context of such narratives that, according to Goldie, they are intelligible to us. Part of the way in which narratives play such role has to do with the fact that they connect the elements of the experience–i.e., various feelings, thoughts, perceptions etc.–to each other as manifestations of some single emotional 'state', e.g., a person's 'jealousy' of their beloved. To illustrate, consider the description Goldie offers of an instance of jealousy.

> You are jealous because you think that she has run off with someone else. You cannot sleep: your heart and mind are racing all night. While you are getting dressed in the morning you cannot help imagining

> them together, talking and joking about you perhaps, and you are unable to keep your mind on anything else. On the way to work, you see another couple in the distance, one of whom looks just like her, and you practically faint, frozen to the spot in terror. Later in the day, you are preoccupied with work for a while, and then suddenly, like a blow to the body, you see on your desk something of hers which triggers your feelings again, and you think 'If I'm not able to talk to her now then I don't know what I'll do'. The next minute your jealousy takes another turn, and you hope you never see her again; the telephone rings and the thought that it might be her fills you with dread. (Goldie 2000, 14)

Let's assume such a complex of feelings, thoughts, imagining, memories, and perceptions, as it is described, is one single emotional particular. Notice that for Goldie, such complex would at best be *part* of an emotional particular: it is an open-ended emotional experience, which might possibly continue for any period of time comprised in the life of the subject.[48] For present purposes, however, we can afford to simplify Goldie's account and treat the complex of episodes described in the above passage as *one* bounded emotional particular, with an onset and an end point. So, on an account *à la* Goldie, such complex qualifies as an emotional particular in virtue of the particular narrative structure in which it is embedded, a narrative which makes it intelligible from the point of view of the subject as a unitary experience–in this case, the narrative summed up by the phrase: "you are jealous because you think that she has run off with someone else." (ibid.). Now, the first thing we should notice about such emotional particulars is that they must be able to 'survive' interruptions:[49] since they owe their unity to narratives of parts of individuals' lives that may well unfold for years, they may also, in principle, unfold for years (op. cit., 16), an order of time which is not compatible with continued emotional experience.[50] The stretch of emotional experience described in the above quote lasts itself long enough to suggest that it plausibly spanned interruptions. It is

[48] See Goldie (2000, 12-16, 141-175).
[49] Goldie (2012, ch.3) explicitly endorses this view and articulates it helping himself of Steward's (2013) concept of an *individual* ongoing process enjoying "modal robustness in virtue of form".
[50] See Mulligan & Scherer (2012, 347): "[t]he relevant maximal period of time [emotion episodes may not outlast] seems to be the period in between two periods of dreamless sleep."

said, in fact, to include periods during which the jealous person is concerned with things unrelated to the relevant narrative until his feelings are *newly* triggered–cf. "you are preoccupied with work for a while and then suddenly […] you see on your desk something of hers which triggers your feelings *again*." One way to refuse to treat these interruptions in engagement as interruptions of the jealous person's emotional experience, and one to which Goldie would resort, is to insist on the fact that the emotion may continue to unfold in the form of 'mood' until it may eventually resurface as full-blown emotion.[51] Bur the vast majority of accounts of emotion agreeing with Goldie's on the processive character of the phenomenon–e.g., appraisal process theories–would still count as many emotional particulars as interruptions in engagement. Scherer (1987), for instance, explicitly says that on his account, as an account of 'full-blown' emotion, the target phenomenon

> [I]s assumed to peter out as the subsystem processes begin to desynchronize and to return to an independent mode of functioning. It is no longer the case that all subsystems are involved when the information processing system has turned to the evaluation of a new, unrelated event, and/or the support and action systems have returned to baseline, and/or the executive subsystem has ceased to process action tendencies related to the original event, and/or the monitor subsystem has directed the focus of attention elsewhere" (8).

Secondly, we should notice that Goldie's particulars span many transient emotion attitudes. If we identify instead our emotional particulars in reference to thick transient emotion attitudes, we may count as many emotional particulars as the many 'turns' taken by the subject's jealousy that are discernible in Goldie's description. For instance, we may delimit an episode of 'jealousy' marked by an evaluation of the situation as involving the loss of the beloved to some rival, followed by an episode of 'terror'–we may imagine, plausibly, that the

[51] See Goldie (2000, 145-148).

prospect of seeing the rival and the beloved together is appraised as a threat in virtue of its likelihood to produce emotional suffering–followed by what seems to be an episode of 'desiring', or 'longing', a bout of 'anger', and, finally, 'dread.' In support of this manner of carving up Goldie's emotional particular, one might point out that one can identify distinct elicitors for each of the smaller emotion episodes: Goldie mentions *imagining* the beloved and the rival being together, *seeing* a couple in the distance and then one of the beloved's belongings on the desk, and *hearing* the phone ringing. Finally, although Goldie's example of jealousy is silent or not detailed enough on the person's physical state, we may add that focusing on somatic criterial features–e.g., facial expressions, patterns of somatic feelings, levels of arousal etc.–would yield even more, and smaller, emotional particulars, lasting on the order of seconds or minutes.[52]

This quick illustration should suffice to support the claim that depending on which feature(s) of emotion we privilege in defining the concept of 'an' emotion we may end up identifying emotional particulars which partially overlap. A natural explanation of this is that different 'features' of emotion *are* distinct emotional components, and such emotional components, though sometimes synchronised, each have their own temporal trajectory. Robinson (2005), among others,[53] emphasises how widespread such *de-synchronisations* are. In the quote below, she implicitly suggests how they may underlie divergent intuitions regarding emotional persistence or the boundaries of the experience.

> [I]t is noteworthy […] that however quickly a sequence of appraisals takes place, the physiological symptoms of emotion *may last longer*. Suppose I affectively appraise something as a threat but then cognition tells me I am in no danger. Even after I realize I am safe, adrenaline continues to course through my blood for a while. Indeed, hormonal reactions to threat may last hours and even days. By

[52] Cf. Ekman (1972; 1994); Prinz (2004); Scarantino (2014).
[53] Deonna & Teroni (2012, 88).

> the same token, that's why I can still 'be angry' after discovering I haven't been wronged: my facial expression and physiological reactions may still keep going after the relevant evaluation has been rejected [… or else they] may disappear even as the evaluation of a situation persists. (op. cit., 78-79).

Such striking mismatches in time-course make it highly unlikely that definitions appealing to features tracking different emotional components will yield the same emotional particulars. Of course, this cannot explain all disagreement in the emotion literature, for one may also provide different definitions of the same kind of emotional particular. For example, approaches privileging the thick attitudinal-evaluative component may diverge importantly on how they understand evaluations–e.g., as felt bodily attitudes (Deonna & Teroni 2012) perceptions of value (Tappolet 2016), 'feelings-towards-value' (Mitchells 2021b), affective responses to intuitions of value (Mulligan 2008, Müller 2019) etc.–and such disagreements are very often substantial. However, at least some disagreement is more deeply ontological in the sense I have tried to clarify.

What I wish to suggest now is that this partial diagnosis of the lack of consensus on how to define 'emotion' provides us with initial reasons to turn to emotion episodes as attractive candidate emotional particulars. Many common definitions of the concept of 'an' emotion seem in fact to presuppose a preliminary operation of abstraction by means of which *the* emotion is identified with some component or other of what one may treat as unitary complex episodes, namely the totality of the changes that we undergo in reaction to cognised emotional situations. Due to the synchronic and diachronic complexity of emotion episodes, such abstractions are likely to yield a number of widely diverging definitions. Because of the synchronic complexity of emotion episodes, emotion theories may widely diverge as regards the change or sub-set of changes they identify as the emotion–having done this, as Soteriou (2018) puts it, they typically "consign the remainder to one or more of the following categories: cause of the emotion, effect of the emotion, occurrence which is a manifestation

of the emotion" (71). Since distinct components possess different temporal trajectories, emotion theories may also diverge in how they trace the temporal boundaries of their emotional particulars. Assuming all this is convincing, if we have some grasp on the concept of 'emotion episodes' independently of assumptions regarding what the *essential* feature of emotion is, why not remain neutral on which component, or set of components, is to be identified as *the* emotion, and attempt instead to provide an analysis of the relations holding between the various changes we undergo during emotion episodes? With such an analysis in hand, one may then go back to the question of the essential features of emotion better equipped to approach it–e.g., having identified the relations of causal dependence among emotional components. Moreover, it might turn out that no single component of emotion episodes is essential to it. On the one hand, each component, arguably, may also occur in isolation. On the other hand, the *essential* feature of emotions as kinds of syndrome, may well consist in the *simultaneous* occurrence of various changes and their distinctively emotional interactions. The availability of an approach focusing on the nature of relations between components, as Soteriou (2018) suggests, may thus help "temper what may be false hopes and expectations in the study of the emotions," namely, "that we should be able to pare down the complex constellation of mental and bodily phenomena associated with any given emotion until we have identified the single mental event or state which is the emotion" (89).

Before moving onto the next section, I want to make a further suggestion, namely that failure to view this sort of approach as an alternative is one of the causes of the tendency to conceive of 'emotion' as an essentially theoretical construct. A version of this claim is articulated by Wollheim in the diagnosis he provides of the lack of shared starting-points and intersections amongst theories of emotion. Wollheim says that as one sets out to study the emotions one may look for two kinds of 'assurance' of the fact that these phenomena constitute a unitary object of inquiry, namely 'prior assurance' and 'posterior assurance' (Preface, xii). Prior assurance is assurance that emotions are *already* individuated, pretheoretically, and one may

try to obtain it, Wollheim suggests, in either of two ways: by appealing to the language of emotion or by referring to existing emotion theories.[54] He states that both attempts are destined to fail. The first option is dismissed by casting doubts on the idea that specific emotion concepts are determinates of the determinable 'emotion'.[55] As regards the theories, as we have seen, he takes it that it is just not clear that by the word 'emotion' they mean the same exact entities. He then goes on to suggest that posterior assurance is the only assurance that one should expect to obtain. Now, Wollheim's discussion of the nature of posterior assurance reflects an attitude towards the question of the unity of emotion that, I will suggest, we can oppose by putting forward an account of emotion episodes. Wollheim describes how we may obtain posterior assurance as follows:

> First, we survey what we pretheoretically assume to be the emotions. Next we draw up an account which fits the material as best we can manage. […] We can think of the account as giving us posterior assurance if (one) the fit is reasonably neat, (two) the account coheres with such intuitions as we have about the emotions, (three) any exceptions to the account that we observe fall into one or other of two categories: either, on further consideration, their claim to be cases of emotion comes to seem less plausible, or their claim continues to hold up, but they are evidently emotions of a more complex sort so that it is only to be expected that, if they are to be accommodated, it will be at a later stage of the account. […] It cannot be denied that, in any philosophical account of the emotions that is rich enough, there will be an element of *stipulation*. At the point at which this occurs, the inquiry turns back on itself. *Strictly speaking, we can no longer claim to be setting out the truth about the emotions as we find them*: *from this point onwards, the emotions are whatever our account is true of*. (ivi, 16-17, italics added)

[54] It may sound odd that theories of emotion may provide prior assurance, but the idea here seems to be that widely shared "starting-points" and "intersections" amongst the existing theories would provide evidence of the pre-theoretical individuation of the phenomenon (cf. op. cit., Preface xiii).
[55] Deigh (1994), as we have seen, makes a very similar remark in terms of *species* and *genus*.

Wollheim seems to think that one should not expect to be able to individuate the *explananda* of an account of emotion prior to having completely articulated one's definition of the concept of 'emotion', and his reason to think so seems linked to the idea that the concept of 'emotion' is at least partly *stipulative* or technical: it is one of those concepts whose boundaries are fixed by the scope of the theories within which they are constructed. As such it is a concept that cannot be fruitfully analysed if not contextualised case by case within the theory which makes use of it (cf., e.g., Izard 2010; Russell 2012; Scarantino 2012). Now, it is important to bear in mind that Wollheim is assuming that the *explananda* of an account of emotion are the various specific emotions, that is, phenomena identified in reference to emotion attitudes, and I agree that there is likely to be an element of stipulation in all accounts of emotion the aim of which is to account for all the referents of our disparate emotional categories.[56] Relatedly, some have argued that, in so far as we seek some minimal consensus on the issue of defining 'emotion', we should ignore our *folk-psychological* conception of such phenomenon and replace it with emotional constructs that are suitable for scientific investigation–for example, primitive emotional components, or some suitable sub-class of emotional phenomena.[57] My proposal will be that we shall be able provide an account of "the emotions as we find them" if we abandon certain widespread theoretical assumptions which prevent us from treating the emotion episodes we undergo as emotional particulars, and most notably, the assumption that the natural *explananda* of an account of emotion are the various specific emotions. I will now turn to the task of characterizing emotion episodes independently of emotional categories. I will start with a very abstract characterisation, and I will then build on it by contrasting emotion episodes with other affective episodes.

[56] Wollheim's explicitly identifies the emotions with enduring emotion attitudes–including 'love' and 'hate', which many treat as 'sentiments.' Here is one of the elements of stipulation of his account. The account, however, is not one of my *polemical* targets: it explicitly concerns dispositions, as opposed to episodes of emotion identified in reference to continual attitude retention.

[57] Cf., respectively, Russell's (1999; 2003) 'core' affect or Griffiths & Scarantino's (2011) 'biological emotions'.

2 The naïve ontology of emotion

Let me start with the idea that emotion is episodic, that is, it *comes* in discrete episodes or temporally bounded occurrents. This very basic idea seems to be at the core of our naive ontology of emotion: we speak of *fits* or *bouts* or *bursts* of emotion,[58] and when we recount instances of emotion, we seem to treat them as psychological continuities with *given* duration– i.e., onset and offset. In other words, we seem to think and speak of emotion in ways which implicitly recognise the pertinence of questions such as "How long did the emotion last?", "When did it begin?", "When did it terminate?" and the like, Now, such bounded psychological continuities are not the sorts of continuities that may steadily unfold over time. If we think of emotion as episodic, as we will see in more detail later on, it is in part because emotion has intensity, and more precisely, it has intensity which tends to vary over time: emotion episodes are momentary global *perturbation*s across affect, somatic feelings, attention, motivation etc.[59] As the emotion episodes we undergo persist and unfold over time, emotional intensity outlines their distinctively dynamic trajectories, namely a course of development which goes through the stages (or phases) of rise, peak(s), plateau(s) and finally subsidence. This element of the concept of emotion 'episodes' is also at the core of our naïve ontology of emotion: we think and speak of emotions as slowly or quickly building up to their peaks, as lingering on and as gradually or suddenly subsiding.[60] A further, closely related element of our naïve ontology of emotion is the idea that emotions are *temporal* entities which *run their course*. I shall take this point to be especially significant. It conveys the idea that

[58] Cf. Scherer (2014) on 'affect bursts.'

[59] Understanding occurrent emotions as momentary ('disorganising') disturbances or *perturbations* of affect and therefore cognition, was the orthodoxy in psychology up until the end of the second half of the last century– cf. Leeper (1949); Young (1949). Cf. Ryle's (1949, §4.3) discussion of 'agitations'; Peters & Mace (1961).

[60] The idea that we naively attribute to emotion episodes such lifecycles is supported by diary studies discussed in Frijda *et al.* (1991) and Frijda (2007, 179-185). The view that emotion episodes possess such lifecycles is emphasised by Bain (1865): "an emotional wave once roused *tends* to continue for a certain length of time. […T]he physical tremor and the mental consciousness pass through successive stages of rise, culmination, and subsidence" (120, italics added).

emotion episodes possess what we may call 'lifecycles:' they have *natural* temporal trajectories and *natural* temporal boundaries, since, having been initiated, they tend to naturally terminate once they have run their course.[61] In what follows, I will explain what I mean exactly by 'natural' temporal boundaries since such qualification is crucial to appreciate the significance of the claim that emotion is episodic. To do so it may be helpful to compare the concept of an 'episode' as it applies to emotion with the same concept as it applies to other types of psychological continuities such as, e.g., perceptual experience.

First of all, we may note that the concept of episodes of perception does not seem to be part of our naïve ontology of perception in the same way in which the concept of episodes of emotion is part of our naïve ontology of emotion. To wit, questions regarding the duration and temporal boundaries of (unqualified) perceptual experience, like 'how long did it last?' or 'when did it begin and terminate?' do not seem as common or as pertinent as questions regarding duration and temporal boundaries of emotional experience. This point, however, cannot be pressed too far, for such questions are clearly intelligible and they are often raised, for example, in philosophical discussions of perceptual experience: perceptual experience is susceptible of being delimited in various manners so as to identify discrete episodes of it. For example, it can be delimited in reference to full periods of consciousness (cf., Tye 2003), in reference to the various objects which, over each period of consciousness, successively happen to occupy the perceiver's attention, or in reference to the active sensorial attendings–e.g., listening or looking–the perceiver is engaged in. We may speak, for example, of episodes of 'listening (or looking) to this sound (or at this object)' or of 'looking out of the window.' We can also imagine situations in which one delimits stretches of perceptual experience in reference to temporal intervals of arbitrary duration and treat them as separate episodes of perceptual experience–e.g., one may want to measure how well the subjects of an experiment

[61] As we will see in the next chapter, the idea that emotion episodes run their course is related to their being kinds of *reaction* and their enjoying the kind of passivity that this notion implies.

perform in timed tasks. Now, if this were what is meant by saying that emotion is 'episodic', then, this preliminary ontological assumption would imply little more than emotion, just like perceptual experience–and indeed any occurrent with temporal extension–is susceptible of being *carved* up in many ways depending on one's interests (cf., Hornsby 2012; Steward 2013). It is true that emotion, just like perception, is susceptible of being delimited in many ways. Indeed, identifying emotion episodes in reference to transient and standing emotion attitudes does just that: depending on how one then specifies the objects or contents of these states one can delimit discrete stretches of an otherwise unbroken continuity and identify distinct episodes of emotion as episodes *of* 'fear', 'shame', 'gratitude', 'resentment', 'love', 'hate' etc–where 'of' may have various meanings. Nonetheless, unlike perceptual experience, emotional experience is also *already* delimited, i.e., independently of reference to any relational property such as its being directed towards worldly objects. That this is the case seems to be confirmed by the fact that *unqualified* emotional ascriptions, i.e., ascriptions reporting the *mere* occurrence of emotional experience, are themselves informative–consider, for instance, 'to be moved' in one of its meanings, as in "I liked the play, but I wasn't moved."[62] By contrast, an ascription reporting the mere occurrence of perceptual experience, absent further specification in terms of objects or contents of perception is not usually informative–barring, of course, those cases in which the capacity for or the very possibility of perceptual experience is being questioned–consider, e.g., "can you perceive *anything* in a 'float-tank'?". These intuitive contrasts between emotion and perceptual experience confirm that saying that emotion is episodic does not *just* mean that one can delimit stretches of emotion, but that emotion is somehow inherently episodic: it 'comes' in episodes independently of the way in which one may then choose to (further) delimit it in reference to its objects or contents or any other relational criterion. This idea proves its significance if we contrast emotion and *mood* as 'mere' emotional affect. According to several theories, mood is in fact ubiquitous: we are always in some mood or

[62] Cf. Deonna & Teroni (2012, 4) and Claparède (1928).

other or some default affective condition. Watson (2000, 13), for example, argues that "[w]aking consciousness is experienced as a continuous *stream of affect*, such that people are always experiencing some type of mood". This is why we can always check in with ourselves and see how we feel, even though our normal or default affective condition is distinctively attentively 'recessive.'[63] Emotion, by contrast, is sporadic or occasional. Indeed, we tend to think of it as an exceptional occurrence: a momentary *perturbation* of our default affective state, that, as such, enjoys relative salience.

So far, I have clarified in what sense the ontological assumption that emotion is episodic and dynamic is part of our naïve ontology of emotion, and that the source of this idea is that such episodes are discernible in experience as bouts of affect: temporally bounded alterations of our default affective condition, which, having been *initiated*, tend then to run their course. While the occurrence of bouts of affect is necessary for the occurrence of emotion episodes, it is not sufficient, for other kinds of episode in the stream of affect are also individuated pretheoretically, e.g., bouts of 'elation', 'anxiety', 'depression', 'melancholia', 'restlessness' etc. Such affective episodes are often called 'moods.' To distinguish this use of the term 'mood' from the one which signifies the stream of affect in its continuity, I will sometimes use the phrase '*specific* moods', or 'specific mood changes.' The issue we need to address to secure the results of the above discussion is whether there is some clear and non-graded distinction between emotion episodes and specific mood changes.

3 Emotion episodes and specific mood changes

In the context of discussions of the differences between occurrent emotions and moods, three distinct respects of contrast are commonly mentioned: *duration*, influence on *attention*

[63] On this see, e.g., Dainton's (2000, 31-32) discussion of mood as an element of the phenomenal background. Cf. Russell & Barrett's (1999, 806) notion of 'core' affect and Tappolet (2018, 169).

and strength and *specificity* of motivating dimension.[64] Occurrent emotions are short-lasting–on the order of seconds or minutes–highly attention consuming, and they involve impulses to engage in specific actions–towards specific 'objects'–which enjoy control precedence. Specific mood changes, by contrast, are relatively long-lasting, have diffuse *biasing* influences on attention and involve a "low ongoing readiness" (Frijda 2007, 51) to act in certain *manners*–for example, 'angrily', 'wistfully', 'restlessly' etc.–which doesn't interrupt ongoing activities. Now, such features are not well suited to yield the principled distinction we need, for they only afford graded, dimensional distinctions, compatible with the view that emotion episodes and specific mood changes are *varieties* of one and the same kind.[65] This view receives support from the fact that occurrent emotions and specific mood changes sometimes come in pairs, and members of each pair have similar phenomenal characteristics and similar cognitive (evaluative) *significance*:[66] For example, fear, anger and sadness are rather similar, respectively, to anxiety, irritability and depression, and this both in the way they feel and in their being associated to thoughts intelligibly connected to analogous evaluations–e.g., episodes of 'fear' and anxious moods have both to do with 'threat.' Moreover, members of each pair often follow one another in the stream of affect. For example, anger tends to leave us more or less irritable, and irritability tends to lead to bouts of anger. If one can only find graded criteria, it may be tempting to *blur* the boundaries demarcating emotion episodes and specific moods in the stream of affect–as Goldie does to allow his emotional particulars to span interruptions in engagement–that which would threaten the episodic character of emotion. To secure the distinction we must thus find some categorial difference, and, ideally, one that may allow us to do justice to the striking similarities I have mentioned.

Some have argued that specific mood changes differ from emotional particulars in that they do not have proper aboutness or intentionality, that is, the property of being directed towards

[64] See, e.g., Frijda (1994); Morris (1989, ch.2); Siemer (2009).
[65] See, e.g., DeLancey (2006). Cf. Tappolet (2018) for discussion and further references.
[66] See, e.g., Goldie (2000, ch.7); Mitchells (2021); Tappolet (2018, 177-180).

worldly objects (Deonna & Teroni 2012). This view is promising, but it is controversial.[67] There are many specific moods for which we have names, and it has been argued that they differ from each other precisely in the "manners in which the world, and the subject's place in the world, are *apprehended*" while undergoing them (Crane 1998, 242, italics added), that is, in their intentionality. If we are persuaded by this line of thought and chose to grant some kind of intentionality to specific moods, then, due to the analogies in cognitive significance between emotion and mood in each 'pair', the distinction once again turns out to be graded. Goldie (2000), for example, contends precisely that what distinguishes emotional experience from mood experience is just that "emotions have *more specific objects*. The distinction is thus a matter of degree." (17).

Assessing the many views articulating the distinction between occurrent emotions and specific mood changes is beyond the scope of this thesis. Moreover, I wish to remain neutral on the question of whether specific mood changes possess some kind of intentionality, since I think that we can secure the distinction at issue without committing ourselves to an answer to such question. Let me start by noting that most proposals on the market seem to share the same kind of approach: they start with two lists of categories signifying, respectively, occurrent emotions and specific mood changes, viewed as isolated, independent phenomena, and enumerate the differences. The previous discussion of the episodic character of emotion suggests, however, a different approach, which promises to preserve the differences while accounting for the striking similarities: if emotion episodes at least partly consist in a bounded alteration in the stream of affect, then, they must involve specific mood changes as their components. If this is so, to compare emotion episodes and specific moods is not to oppose two similar phenomena on a par with each other; rather, it is to oppose a *whole* with one of its *parts* in isolation. This idea fits the view that specific mood changes tend to occur–or more

[67] See, e.g., Crane (1998); Goldie (2000: ch.7); Prinz (2004: ch.8).

precisely, to come into view as such–typically in the *aftermath* of emotion. In Ruckmick's (1936) words, a mood is "a degenerated or decayed […] emotion, one that […] follows in the wake of an emotion, like an attenuated tail or stem" (72). That emotion episodes and specific moods are taking place simultaneously does not imply this sort of account, for the idea is also captured by an account treating emotion episodes and specific mood changes as two separate affective processes that are triggered simultaneously and that operate in parallel fashion.[68] However, if our basic units of analysis are emotion episodes as emotional wholes or unities including 'mere' emotional affect, nothing forces us to opt for an account in terms of two parallel processes such that when the emotion process is over the (milder) mood process appears. Furthermore, as we will see, it is an attractive character of the view that emotion episodes involve specific mood changes as one of their components that it allows us to ascribe to the emotion features that many accounts, in their attempt to clearly differentiate emotional particulars from specific moods, reserve to the latter. In fact, such features, especially the diffuse biasing influence specific mood changes seem to have on our interests and attention, will turn out to be crucial to explain the persistence of emotion episodes. The proposal is thus that anger, for example, involves *from the start* irritation: typically, the angrier we are, the more irritated (and the more irritable) we become, and when our anger subsides, if the episode was intense enough, it may take some time for us to return to our previous affective condition. This idea can explain common interactions between emotions episodes and specific mood changes: specific moods facilitate the elicitation of the episodes of the kinds of emotion with which they come in pairs–e.g., being irritated lowers the threshold for anger–and, conversely, episodes of given kinds of emotion leave behind them corresponding specific moods–e.g., 'being angry' leaves us irritated for some time (e.g.,

[68] See Morris (1989): "Clark and Isen (1982) […] made the observation that mood and emotion can be instigated by the very same event because *the two processes can operate in parallel.* However, in this case one may not be able to detect the presence of the mood because the emotion has the more powerful and dramatic effects. As the emotion wanes, the influence of the mood may appear. (12, italics added)

Ekman 1994, Frijda 1994, Lazarus 1994). This will gain further support from the analysis of the causal structure of emotion episodes I will offer in the next chapter; the analysis will in fact reveal that these recurrent transitions from specific moods to emotion episodes and *vice versa* are explained by the same interplay (of affect, interest, and sustained attention) that underlies the way in which emotion episodes *unfold* over time. For now, I will suppose that this picture is correct: our emotion episodes differ from specific moods as wholes differ from their parts. I will now turn to the issue of the boundaries separating emotion episodes from the specific moods they leave behind. Discussing this issue will allow me to put forward the suggestion that emotion episodes are individuated in reference to attention. As anticipated, we will thus end up with a dual, *conjunctive* criterion for emotion episodes: they are bouts of affect *and* episodes of sustained attentional engagement.

4 Pears on the link between the duration of emotion and attention

The general idea I wish to explore is that the way in which our attention is deployed in emotion is different from the way in which our attention is deployed during specific moods and that we can use this momentary modulation of attention to draw temporal boundaries demarcating emotion episodes from specific mood changes in isolation. This idea is not new; it is explicitly put forward by Pears (1962) in his discussion of the disanalogy between ascriptions of 'momentary' *vs* 'persistent' psychological reactions–viz., affective episodes–with respect to the possibility of errors through object-misidentification. Pears' aim is to understand why self-ascriptions of affective episodes like 'amusement' seem to be immune to such errors, whereas persistent affective episodes like depression are not. I will not need to get into the details of Pears' discussion. What interests me is just the argument Pears develops to establish the relevant disanalogy. Pear's central proposal is that such disanalogy is connected to the fact that while we always identify momentary affective episodes in reference to some object, we can sometimes identify persistent affective episodes without

reference to any object–i.e., as *free-floating* feelings *vs* feelings that are reactions to an object. Pears introduces the disanalogy in the passage I quote below.

> Some feelings or reactions are momentary, like amusement, and they tend to be individuated together with their objects. For there is simply no question whether the amusement that someone is feeling at this particular moment might be identical with an earlier feeling of amusement which had, or seemed to have, a different object. Contrast the other end of the spectrum, when a feeling like depression or anxiety is persistent, so that the person himself at the time when he is feeling it can *sometimes individuate it in a minimal way that involves no reference to its object (i.e., it can be free-floating)*. In such cases one can ask oneself questions of identity that are in many ways analogous to questions of identity about physical objects. For instance, is the anxiety that I am feeling today still yesterday's anxiety about a task that I rashly undertook? And the answer to this question can never be tested. (ibid., 74, italics added)

The analysis of this contrast that Pears offers leads him to conclude that relations between affective episodes and 'objects'–i.e., what they *seem* to the subject to be about–are causally mediated by attention: if an emotion or mood is to have an item as its object, it must be *caused* by the presence of that item to the mind of the subject, i.e., by the subject's attending to that item.[69] His argument turns on the following observation.

> How would it be established that someone had mistakenly identified the object of his depressions? […] First, *he might cease to feel depressed about the stated object as soon as his attention was drawn to the rival object, about which he would immediately begin to feel depressed*. Of course, in this situation he might say that one bout of depression had been replaced by another. But he would not have to say this […]. For in certain circumstances he would use a *different criterion of identity* for bouts of depression and he would say that he

[69] Cf. Wilson (1972) for one development of this proposal. Some philosophers deny that the relation between emotion and object-directed thought is even in part causal arguing that it is logical, or normative, or constitutive (cf. Thalberg (1964); Kenny (1963)). The typical argument is an argument exploiting the non-contingency of the relation between emotion and object. Wilson (1972) argues that the argument from non-contingency fails, roughly, since there is no real opposition between causality and such kinds of necessity. For more recent discussion of these issues, see Teroni (2017).

> had discovered that the bout of depression which according to the first statement, had one object, really had another. (Pears, 1962, 64; italics added)

The idea here is that what makes room for errors through object-misidentification in cases of depression is that although a criterion of identity for bouts of depression in terms of *causal* relations between these bouts of affect and attention to an item is available–i.e., shifts from one item to the next may be taken to mark the beginning of 'new' bouts of affect–crucially, this identity criterion is only optional: the fact that the subject may start to be depressed about a new object as soon as their attention is drawn to it might be taken as well to reveal that the new object is what one was depressed about all along.

This optionality highlights that some other criterion of identity is available to the subject; this other criterion is not clearly defined by Pears–I will attempt to specify it in due course–but it is linked to the fact that feelings of depression are sometimes considered 'free-floating.' What matters is that, in any case, the contrast with affective episodes fixing the other end of the duration spectrum, such as, e.g., startle, surprise or amusement is rather sharp. Suppose one ceased to *feel* amused by a joke as soon as one's attention were drawn to some rival joke, about which one would immediately feel amusement. In such cases, it seems that one would readily accept that a *new* bout of amusement replaced the old one. The rival identity criterion, according to which there is but one bout of affect, which once seemed to have an object, but then turned out to have another–cf. e.g., 'that's what I was *really* amused about all along'–is not available for amusement in the way in which it is for depression. Therefore, *momentary* affective episodes, unlike persistent affective episodes, are identified in reference to one *single* criterion, that is, the presence of an object to the mind of the subject as ensured by attention.

What remains unexplained is the nature of the relation between the duration of the episodes and their different object-relations. Is this a *brute* psychological fact or can it be analysed?

Pears' claim is that the more long-lasting or persistent the episodes are, the *looser* their link with objects becomes–i.e., the more likely it is that such episodes be identified as *free-floating* feelings as opposed to reactions that are triggered anew at each shift. The idea here is that because of their longer duration, persistent affective episodes get more chances to "collect other things […] the person would then be equally ready to identify as […] objects" (75). This suggests that Pears would take emotion episodes and specific moods to differ in degree. The view seems implied by his speaking of 'amusement' and 'depression' as fixing opposite extremes of a spectrum of duration, whereas 'fear', 'anger,' 'shame,' etc.–affective episodes we would readily classify as emotional particulars–lie toward the center. As regards the question of what explains the duration of different episodes, he seems to think that this feature is a brute (species-specific) psychological fact (op. cit., 74): amusement is very short; depression lasts much longer; fear, anger, shame etc. lie toward the middle. This may seem intuitively appealing. As we will see in the next chapter, it is not uncommon to see the duration of emotional reactions as species-specific and such view receives some support from empirical evidence showing that different emotions differ in their average durations. For example, joy and sadness tend to last longer than anger, and fear rarely exceeds one hour (Scherer *et al.* 1986; Scherer & Wallbott 1994). This psychological fact, however, still calls for explanation. A plausible explanation is that different emotions involve different constraints on the kinds of object that may elicit them, and this, in turn, affects average duration. Verduyn *et. al.* (2009, 86) for example, suggests that since fear is "typically a future-oriented emotion, whereas sadness and anger are past-oriented emotions […] fear may last until the fear-arousing event ends, whereas dwelling on the past can be endless". Now, we may notice that this sort of explanation, by identifying mental activities, such as 'dwelling on the past', as *mediating* variables, makes the link between kinds of emotion and the duration of instances of such kinds rather loose: whether or not one will engage in mental activity that may prolong the episodes' duration will depend on several contingencies. Frijda (2007, 191), in assessing

the empirical research on emotion duration, suggests that species-specific average duration is "largely due to the ecology of the emotions, rather than […] kinds of emotion as such. Joy and sadness, in these studies, mostly concerned relationships, that usually extend in time; and fear mostly concerned […] transient events". Generalising Frijda's observation, one may suggest that the typical objects of a person's emotional reactions of a given kind will depend on the sort of life that that person lives, and that this is an extremely pervasive ecological constraint on the duration of that person's reactions. In a war zone, for example, episodes of fear and anger are likely to exceed reported *average* duration for such kinds of emotion, and, intuitively, they are likely to persist longer than both joy and sadness. Besides the variable features of the typical emotion-eliciting situations for given kinds of emotion in different contexts, the duration of emotion episodes plausibly depends on many factors closely related to the likelihood of the occurrence of mental activity that may prolong emotion, for instance, the capacity of the subjects to habituate through exposure, the intensity of the reaction, availability of suitable distractors etc. Nothing prevents, in principle, episodes of fear or anger from lasting minutes, or an hour, and episodes of joy or sadness from waning within seconds (cf. Frijda 2007, 185-188). Now, the idea that duration may depends on cognitive activities such as, e.g., 'ruminating' or 'anticipating' suggests that one may retain Pear's suggestion regarding the existence of the link between the duration of affective episodes and the nature of the episodes' object-relation while reversing the proposed order of explanation. Pears' mistake, I am suggesting, is to attempt to account for the nature of the object-relation distinctive of different affective episodes by appealing to their duration as a species-specific fact, and not the other way around. The proposal I will try to articulate is that it is because occurrent emotions and the specific moods with which they come in pairs relate to their objects through attention in different manners that their typical duration is different.

5 Emotion, attention and individuation

We may try to sharpen the contrast between specific emotions and specific mood changes by considering those cases in which we think we know exactly what elicited a mood change–for example, what made us feel depressed–and we would not be inclined to doubt that even if we started feeling depressed about a rival object as soon as we thought about that object. Suppose that I learn that a paper I have submitted has been rejected. Upon receiving such bad news, I start feeling depressed, and then continue to feel depressed throughout the week. Suppose that this involves having frequent thoughts about the paper rejection and the occurrence of episodes of 'sadness' I identify as caused by the occurrence of the object-directed thoughts.[70] Plausibly, my being depressed involves as well the occurrence of episodes of sadness towards things by which I wouldn't *normally* be affected, or not as much, and which do not seem to be related to the paper rejection–for example, I may feel sad about the fact that my friend won't have time to grab a drink with me this week. I take it that such cases are the intuitive sources of the view that our emotion episodes and specific mood changes differ in degree. Unlike those cases in which we just find ourselves in specific moods without clear reasons, in this scenario, I am aware of the initiatory cognitive elicitation of my specific mood change–namely, I started feeling depressed *upon* receiving the bad news–and, relatedly, what I'm really depressed about remains transparent to me for the duration of my depression: if I find myself feeling sad *that* my friend will not grab a drink with me, when reminded that this is the case, I am not tempted to suppose that this is what I was *really* depressed about all along. Indeed, I may even know that my feeling depressed about the paper rejection importantly contributes to how easily and how intensely the thought that I will not see my friend is affecting me–perhaps I had been reminded of this fact many times

[70] I will use the term 'sadness', and not 'depression,' for bouts of depression that accompany object-directed thinking, that is cognisance of the *object* my depression is *about*. This is perhaps a bit tendentious, since it is often assumed that sadness is the properly emotional counterpart of the mood 'depression,' but it will make my point a bit clearer; the reader is free to replace 'sadness' with 'depression.'

before the paper rejection, and, while I had been slightly upset each time, I had never become too sad. This does not mean that I am not sad that I will not see my friend; what *really* saddens me, as I am thinking about my friend, is that I will not be able to see him, and while the mood contributes to my sadness, it doesn't, so to say, *steal away* its object: I feel sad about one thing, while being depressed about something unrelated. Now, the fact that moods do not steal the objects of the emotion episodes to the occurrence of which they seem to contribute is still compatible with the idea that emotional and mood-like reactions are *both* identified in reference to what they seem to be about–and, therefore, may arguably differ in degree. It remains correct to say that, throughout the week, I am depressed *about* the fact that my paper was rejected, and one may argue that *this* feeling, as it persists, sometimes focusses into sadness about the paper rejection and sometimes–for example, when I think about, and react emotionally to other objects–it recedes a bit to the *background*– without thereby ceasing to be about what it initially seemed to be about. The idea that our specific moods do not steal away our emotions' objects, however, allows us to see how the sense in which my sadness is about the paper rejection and the sense in which my depression is about the paper rejection are different. Even when I am currently thinking about the paper rejection, and, while doing so, I undergo proper episodes of sadness–that is, bouts of affect I readily identify as caused by my object-directed thinking–I can still say that the occurrence and the intensity of such episodes is partly determined by my being depressed *about* the paper rejection, and this in the same sense, it seems, in which I can say that unrelated episodes of sadness are partly determined by my depression. There is in fact an explanatory relationship between episodes of sadness and persistent depression that is not compatible with the idea that sometimes my persistent depression is *identical* with episodes of sadness while some other times it recedes to the background of my experience. Such explanatory relationship seems to obtain whether or not the object of my sadness is the same as that of my depression, and

this suggests that their being *about* the same thing, when they happen to be so is still crucially different.

The natural suggestion is that by saying that I have been depressed about the paper rejection throughout the week I am saying that *receiving* the bad news, at the given moment that I have, has *made* me depressed for a whole week; in other words, the 'about' adverts primarily to the fact that since and because I received the bad news, I was disposed to look at things in a depressed way for a whole week. The identity of specific mood changes, whether or not we are aware of their objects, is not tied to our being currently cognisant of such objects; rather, the tie to their objects is ensured by the causal and temporal continuity with respect to their initiatory cognitive elicitation, or more precisely, to the elicitation of the emotional reactions of which they are the 'trail.' It is in this sense that specific mood changes *arise* in the aftermath of emotion. This sort of continuity, I suggest, is the *rival* criterion for persistent reactions that Pears has in mind, one that becomes available to subjects of specific mood changes after their first emotional reactions have subsided. Now, since our moods do not require us to be currently cognisant of what they are about, they do not draw upon our attentional resources in the way in which emotion episodes do. Consequently, they are free from constraints related to the limited character of such resources: although they compete for such resources with all other affective and cognitive processes, when they lose the competition, they may recede to the background stream of affect. This is how they get the chance to 'collect' new objects during their course. By contrast, emotion episodes are short-lasting in comparison to the specific moods changes they leave behind because they require our continual occurrent cognizance of emotion-eliciting particulars or situations and are thus subject to the limitation of our attentional resources.

I will now conclude this discussion by going back to the idea that specific mood changes occurring in isolation are the aftermath of emotion episodes and assessing an objection to it.

This idea fits nicely cases such as the one I have described. It is plausible that upon receiving the bad news about my paper I undergo an episode of sadness, which then lasts until I direct my attention elsewhere. At this point, the reaction subsides, leaving me feeling depressed. Now, to fully support this suggestion, we should be able to apply it to those cases in which we are not aware of the initiatory cognitive elicitation of the reaction–or not so keenly aware. One way to do so is made available by Frijda (2007, ch.2). Frijda suggests that most of our everyday emotional reactions are *inner* or 'small' emotions. A small emotion is distinctively *subthreshold* for full-blow emotionality in various regards–e.g., duration, intensity, and impulse to actions. However, as Frijda emphasises, we do feel such reactions and we can often name them–e.g., "I felt sad (or angry, attracted, disgusted) for just a moment". Our small emotions, moreover, often have important cognitive and behavioural consequences, like making us notice new aspects of events or people, and thereby change our attitudes towards them, or prompting us to make crucial decisions, for instance, "subscribing to a particular political party; deciding to donate money to a cause; or to hide a persecuted person, with risk for one's own life." (op. cit., 44). I wish to suggest that one further, *durative* effect small emotions have is to leave us in a mood after their very quick subsidence. Sometimes, for example, the small emotion's elicitor is not itself especially significant; maybe its power to elicit an emotion is inherited from some other thing *via* associations of which we are unaware–e.g., a friend's observation can make us angry for an instant because it reminds us of something that we are angry about–in the single-track sense of 'being angry'–and not because it is itself offensive. In such cases, our attention will not be retained by the elicitor–i.e., the observation–for very long, but we may remain irritated. Alternatively, the elicitor may be too significant, for example, too sad and too awful to think about, and so we 'refuse' to engage with it and divert our attention elsewhere with the same result: we end up irritated, or anxious, or depressed. In all cases, what is more noticeable and cognitively significant is not our 'small' emotion,

but our subsequent mood, and so the latter may seem to arise out nowhere as opposed to arising in the aftermath of our emotional reactions.

I have suggested that we may individuate emotion episodes in reference to attention, for the boundaries of emotion episodes are marked off by an alteration in the stream of affect associated to our being currently cognisant of given objects. Emotion episodes draw upon our attentional resources in a way that does not allow them to outlast the emoting subjects' engagement with given objects. Because specific mood changes belong to emotion episodes as one of their parts, their onset may be said to be marked by the same cognitive elicitation– even though speaking of mood-elicitation will not be pragmatically relevant if the changes do not last in the aftermath of the emotion. The reason why specific mood changes may outlast our emotional reactions is that they can recede to the background stream of affect. Now, the idea that we may individuate emotion episodes in reference to attention will be developed in more detail by articulating the relation between the emoting subjects and the emotional situations in terms of sustained attentional engagement with emotional situations. Attention *to an object* is not quite the right criterion, for we ought to allow for given episodes of emotion to continue as the emoting subject's attention is shifted among various objects. Even without considering multi-attitude emotion episodes, it is clear that if we wish to be able to invoke the relation between emotion and attention to produce a criterion of individuation for emotion episodes, we must identify a more flexible constraint on the occupation or the emoting subjects' attention. Amusement is, in this sense, a 'toy' example.[71] In chapter 1, sustained emotional engagement was characterised as *our being related* to emotional situations as objects of our current interest and sustained attention. The idea was

[71] Cf. Roberts's (2003, §3.3) discussion of emotions as kinds of 'evaluative construals' of emotional situations. "In many emotions we will find that the focus of the construal shifts rapidly and freely so that an element that is at one moment the focus is at another moment one of the terms, with a different element as the focus. Thus a certain fluidity of role among the elements of an emotion is compatible with the emotion's remaining the same emotion despite the shift." (op. cit., 79).

that emoting subjects are always interested in emotional situations, and more specifically, that they are interested in an evaluative dimension of such situations. Such active states of interests drive and govern the emoting subject's attention so that their mental activity is geared towards the exploration of the *value* (or relevance or significance) of the emotional situations–i.e., they thus keep emoting subjects, so to say, on topic, by pre-empting distractions. As we will see in detail in Part II, the fact that sustained emotional engagement is governed by interests specifying conditions for continuation *vs* interruption–i.e., distraction–is what makes it the sort of thing that comes in bounded episodes, and, therefore, something in reference to which one may, in principle, specify precise criteria of emotion individuation. In what follows, I will conclude my discussion of the individuation of emotion episodes by considering the relations between emotion episodes and emotion attitudes and the rationale for the widespread assumption that the latter–or episodes of emotion individuated in reference to them–are *the* basic units of analysis of an account of emotion.

6 Emotion episodes, attitudinal retention and diachronic rationality

Although most theorists working on emotion explicitly recognise that emotion is 'episodic', their chosen emotional particulars are not usually individuated by referring to what I have called *natural* temporal boundaries. They are more commonly individuated as episodes of some kind of emotion. in discussing the episodic character of emotion, I granted that though emotion episodes have natural temporal boundaries, they are susceptible of being delimited in reference to emotion attitudes, viz., in reference to *intentionality* or the cognitive relations emoting subjects are taken to bear to the objects or contents of their emotional experience. It is widely accepted that an account of emotional intentionality must accommodate two different intentional relations: the first relates emoting subjects to particulars, or situations; the second relates them to some thick value, or evaluative property (apparently) instantiated

by particulars or situations (Teroni, 2018). Now, if one individuates emotional experience *via* their evaluative intentionality–i.e., in reference to objects apparently instantiating thick value–one may then identify episodes of emotion as episodes of the kind 'fear', 'shame', 'anger', etc. As seen in chapter 1, there are many ways in which emotion episodes might be taken to be episodes *of* kinds of emotion. According to Goldie or Wollheim, 'emotional particulars' are episodes of kinds of emotion in that they are the manifestation of what I have called enduring attitudes–viz., single-track or multi-track emotional dispositions. On such view, typically, there is little or no interest in tracing temporal boundaries of emotion episodes. Goldie (2012), for example, when he turns to considering emotional experience as it unfolds over time–viz., the sort of processive continuity I have suggested that we may think of as his emotional particulars–is keen to take its boundaries to be fuzzy to the point of including the specific mood changes that emotion episodes leave behind in their aftermath.[72] Furthermore, on this sort of view, as we have seen, 'emotional particulars' are taken to be such that they can survive interruptions, not only in the sense that they may diffuse into specific moods, but also more radically. All that in the life of an individual is an occurrent *manifestation* of enduring emotion attitudes toward given (kinds of) particular or situation–for instance, one's enduring jealousy for one's lover–is part of *one* emotion. In speaking of such view, it is not fitting to speak of 'emotional particulars' since the view does not have as one of its aims *particularising* emotional experience–i.e., drawing temporal boundaries. When Goldie turns to questions of ontology, it is typically to remove boundaries that do not let the phenomenon he is interested in come into view. Goldie's basic units of analysis are especially relevant if one is interested in providing historical or narrative accounts of our occurrent emotional life. Now, I take it that this way of 'delimiting' emotional experience is compatible with the way I have proposed we individuate emotional particulars. Indeed, in so far as the synchronic

[72] Goldie (2000, 145-148) defends such view helping himself of Musil's (1995) interesting discussion of the way in which moods *focus* into emotions and emotions blur out of focus into moods–*The man without qualities* / *From the posthumous papers* (1239-1311). Cf. Mitchell's (2021) recent defence of a version of this view.

complexity of emotion episodes its recognised in both views, Goldie's complements the one I am putting forward, and it may provide the best viewpoint to appreciate the *nested* structure of our emotional life from the most relevant 'time-frame'–for example, by addressing questions regarding the relations between emotion episodes and the complex patterns of emotion episodes to which they often belong.[73]

However, another way of thinking of emotion episodes as episodes *of* some kind of emotion––and, arguably, the orthodox way to do so–does involve commitment to *particularising* emotional experience in a way that is in conflict with the one I am articulating. On this view, emotion episodes are episodes identified in reference to retention of single emotion attitudes understood as 'transient' attitudes–i.e., as *evaluative* attitudes held towards the relevant emotional situations while undergoing emotional experience, and whose continual obtaining over time is (partly) due to the fact that one is currently undergoing emotional experience.[74] What is the rationale for this view? What is there to be gained from this way of particularising emotional experience? Deigh (1994) offers an illuminating discussion of what philosophers must have thought there *was* to be gained from the claim that emotion episodes are episodes *of* some kind of emotion, in this sense, in his reconstruction of what he sees as an historical transition from 'feelings-centred' to 'thought-centred' accounts of emotion. Deigh links the shift to two closely related factors: a shift in the *target* of the accounts from what he calls the 'general' concept of emotion to specific emotion concepts and, second, the gradual demise of the introspectionist methodology in the study of the mind–as championed by James, Broad and Stout, among others–in favour of analytical-linguistic/conceptual methodology.

[73] Cf. Helm (2001) and Na'aman (2019; 2021).

[74] See Soteriou (2013)'s discussion of the concept of an '*occurrent* state' for a fully-fledged articulation of this way of thinking of the relation between what I call 'transient' attitudes–and more generally of any state of mind whose obtaining relative to temporal intervals is essentially connected to the occurrence of given experience over such temporal intervals. See Soteriou (2018) for an application of this ontological framework to emotion. That emotion attitudes are commonly conceived of in ways that are best captured in such terms seems confirmed by the widespread insistence on the claim that the intentionality of emotional experience is inextricably linked to the *experiential* dimension (Deonna & Teroni 2012; Goldie 2000; Montague 2009; 2014; Vanello 2018).

Deigh takes these factors to be related because specific emotion concepts, unlike the *general* concept of an emotion or an emotional experience, were more congenial to the new method of inquiry: in virtue of their close links to thick values, they allowed in fact philosophers to identify emotional particulars suitable for linguistic and conceptual treatment.[75] By contrast, 'empirical' introspective investigations did not have much to reveal about specific emotions, apart from the fact that they each seemed to have distinctive phenomenal characteristics– e.g., Broad's (1930/2013, 228-230) emotional 'tones' or *qualia*. Deigh clearly illustrates this point in the passage I quote below.

> Introspectionism treated emotions as states whose intrinsic properties were all inwardly observable. It treated them, that is, as purely empirical states. To be sure, introspectionists also regarded emotions as signs of other mental states that were their typical causes or effects, and as correlates of neurophysiological states that were their physical underside, basis, or constant companions. But in themselves, as mental phenomena, emotions were not seen by the introspectionists as having any theoretical depth. By contrast, the conceptual and linguistic investigations that replaced introspectionism allowed for an understanding of emotions as having such depth. These investigations concentrated on clarifying the criteria by which we apply the concepts of specific emotions, and one can construe these criteria theoretically. That is, one can take our everyday thought and talk about psychology to contain a theory of the mind whose principles determine these criteria. Indeed, this idea has now become a standard view in philosophical psychology. (op. cit., 833).

To the fact that emotion attitudes lend themselves to linguistic and conceptual treatment we may add that specific episodic emotions are emotional particulars about which one can ask questions regarding 'emotional rationality', and more generally, questions that are pertinent

[75] To illustrate: "[the] refrain typical of philosophers engaged in these investigations went (and still goes) something like this: "There is a logic to the concept of x such that to say that a person feels x towards z implies that the person believes such and such about z." There is a logic, for example, to the concept of pity such that to say that a person feels pity for z implies that the person believes z to be in some distress." (op, cit., 831)

to epistemology and moral philosophy. In virtue of their close connection to thick values, specific episodic emotions provide the emotion theorists with basic units of analysis subject to normative standards: an episode of, say, fear is rational *vs* irrational essentially depending on whether its object instantiates the evaluative properties the attitude 'fearing' assigns to it–e.g., 'fearsomeness' or 'dangerousness'. In general, the emotion attitudes for which we have a name seem subject to normative standards variously understood–for instance, as standards of fittingness, correctness, appropriateness etc.–and whether such standards are actually met is assessable in reference to the objects or contents of the attitudes in question. Now, while there are clear theoretical advantages connected to the assumption that our units of analysis are episodes of emotion identified in reference to emotion attitudes, the view that reaping these benefits requires focussing on the objects of the emotion has recently been attacked for being essentially *synchronic* in ways which obscures the 'diachronic rationality' of some of our emotion attitudes. In a recent series of papers, Na'aman (2019; 2020, 2021) highlights how the rationality of emotion cannot be fully assessed at an instant for various reasons. Firstly, each instance of emotion involves what Helm (2001) calls emotional commitments to feeling other emotions at other times–e.g., 'anxiety' at the prospect of *x* commits us to feeling 'relief' when *x* is over. Secondly, some emotion attitudes may rationally 'self-consume' or 'self-augment' over time–consider 'grief' that ends with some kind of closure, 'resentment' that ends with forgiveness, or 'love' or 'nostalgia' that augment as time goes by. Finally, although Na'aman does not offer a fully-fledged theory of the phenomenon, he emphasises the rationality of what I have called 'sequences of emotions.'

> Rational emotional transitions occur when one attitude is fittingly replaced by another. At first we are excited but then we might grow weary or bored; we view someone as no more than a friend but then fall in love; we are distressed by a new predicament but then reconcile ourselves to it; we can laugh or cry and then calm down. The fittingness of these transitions is not determined merely by the fittingness

> of the former attitude at the time of its occurrence and the fittingness of the latter attitude at the time of its occurrence. Rather, the issue partly concerns how one transitions from one attitude to the next. Furthermore, in these examples the latter attitude is understood and described as emerging from the attitude that preceded it; it cannot be assessed in abstraction from its history. For these reasons, the relation between the agent's attitudes at different times is essential to the fittingness of these emotional transitions. (Na'aman 2019, 17).

Such emotional transitions occur when we react emotionally to unfolding situations. Hence, they are not the sort of emotional transition I'm mainly interested in, for the latter are *proper* attitudinal transitions, or changes in attitude, due to the fact that disparate affective interests or concerns may be at stake in given unchanging emotional situations. But they contribute to illustrating how questions regarding emotional rationality are not limited to the question whether emotion attitudes are fitting with respect to an object at an instant. What I wish to be retained from this quick reference to the literature on diachronic emotional rationality is, first, that whatever the benefits of focusing on episodes of emotion identified in reference to emotion attitudes, these are not lost by rejecting the assumption that our basic units of analysis are specific episodic emotions. Secondly, there seem to be further benefits awaiting the shift from the study of single, 'static' emotions attitudes at an instant to the study of how emotion attitudes rationally *evolve* over time or rationally succeed one another. In other words, one may retain the theoretical advantage due to the historical transition Deigh reconstructs by rejecting the assumption at issue while granting that emotion episodes are meaningfully described in terms of emotion attitudes. An account of emotion episodes treating attitudes as an attitudinal-evaluative component, is not, in principle, at any theoretical disadvantage compared to accounts identifying emotions in reference to continual attitude retention. Concluding, the assumption seems unwarranted. It is the residue of the way in which a given historical transition in the study of emotion came about. The last suggestion I wish to make in this chapter is that multi-attitude emotion episodes are actually the natural *explananda* of

an account of emotion, and, arguably, the most relevant basic units of analysis if one is interested in emotional temporality and 'agentiality'.

Frijda (*et. al.* 1991; 2007) observes that subjects of diary or survey studies, if asked to recount instances of emotions, typically recount 'emotion episodes'–that is, emotional particulars with variable duration and whose 'intensity contours' outline dynamic temporal trajectories.[76] He also emphasises that judging from these reports it is not obvious whether the temporal boundaries of the emotion episodes the subjects of diary or survey studies seem to treat as emotional particulars always, or even typically, correspond to boundaries of single kinds of emotion. Although the emotion episodes are reported in response to an explicit request to describe an instance of some kind of emotion, the subjects also use emotion concepts to identify or describe certain temporal 'stretches' of what they seem to treat as single episodes of emotion. To wit,

> Instances anger are often reported as *including moments* of feeling hurt, hopelessness and fear of retaliation in addition to anger as such and satisfaction about harm done to the opponent and many reports of grief often involve mention of despair, anger and bewilderment, besides sorrow and distress. Reports of jealousy involve distress, anger and sheer anguish besides what subjects label as pangs of jealousy. (Frijda *et al.*, 1991, 188-189).

How are we to interpret this apparent inconsistency? One option is just not to take seriously our use of emotion concepts as referring to emotion attitudes with some psychological *reality*. Robinson's account of the relation between emotion episodes and our emotional categories is an example of this sort of approach.

[76] The subjects in Sonnemanns' study (1991), on which Frijda relies, were asked to draw the intensity contours of the reported instances of kinds of emotion.

> As William James noted, our emotional processes are in constant flux: we can change quite quickly from fearful to cheerful, from anxious to angry. Particular nameable emotional states in this view are typically recognized after the event 'when the emotion has been catalogued in recollection'. So, it is only after the event that we (or our friends) describe a situation as one in which I was sad, angry, or jealous, ashamed, guilty, or merely regretful, or calm, bored, or world-weary. It is by using ordinary emotion words like this that we try to make sense out of and explain our emotional experiences. (Robinson, 2005, 79)

To understand Robinson's contention, it should be noted that although she speaks of shifts from 'being cheerful' to 'being anxious' to 'being angry', she does not take such terms to actually refer to a 'real' component of emotion episodes as she understands such phenomena. In using such terms, she insists, "I am not reporting on the actual sequence of *appraisals and reappraisals* that took place in me. The actual sequence of appraisals that occurs changes very rapidly and I am typically unaware of most of it." (op. cit., 81, italics added). If we consider Robinson's words in light of her explicitly endorsed conception of an emotional 'appraisal' as a non-cognitive, sub-conscious affective appraisal (op. cit., 41-47, cf. Zajonc 1984) it is clear that she does not take emotion episodes to involve emotion attitudes, for she identifies the emotions with essentially sub-personal emotional processes, such that any "cognitive elements of the experience are excluded as not part of the emotion" (Goldie 2012, 58, footnote 4). Now, while it is true that we typically employ emotion concepts when "cataloguing the experience in recollection", in many cases, our emotional experience as we undergo it is an experience of which we have some sort of immediate, or non-reflective, 'categorical awareness' (Lambie & Marcel 2002). Indeed, chronic lack of any form of 'live' categorical awareness is sometimes associated with conditions categorised as pathological or at least clinically problematic–e.g., alexithymia, and more commonly emotional repression

(op. cit., 251-254). Therefore, I do not think that the above reports, and, generally, our use of emotional categories, can be dismissed as Robinson suggests. If we take our ordinary use of emotional categories a bit more seriously, the above reports might be taken to raise a number of questions. First of all, why do the subjects report what may seem to be sequences of emotions as *single* coherent emotion episodes? A natural explanation is that such sequences are felt to have some 'unity' or 'coherence,' that is, that they are experienced as continuities over time whose internal variation marks temporal stretches–e.g., 'phases' or 'moments'–but doesn't constitute interruptions. Frijda proposes two main factors of such unity: (*i*) throughout the episodes the subjects are *affected* by unchanging emotionally significant situations, i.e., they are in an altered affective condition, and (*ii*) throughout the episodes the subjects are interests in and engaged with such situations.[77] On this view, emotion episodes are thus bounded emotional continuities which result from continuous engagement with single emotional situations. Frijda further suggests that we may account for the fact that our emotional categories are also used to refer to certain phases or 'moments' of the episodes in terms of levels of analysis corresponding to a "hierarchical organisation" of emotional processes with different durations (2007, 185). More specifically, he suggests that emotion concepts must be used in the reports both to describe the emotion episodes globally, at the level of what we may call a protracted transaction with emotional situations, whose overall significance the subjects take to be captured by some single emotion concept, and also to more accurately describe variation over time in certain emotional components, such as, e.g., the specific (re)appraisals of the situation and the modes of action readiness – viz., the action tendencies. He then goes on to propose that emotion episodes be viewed as *sequences* of specific emotions. Frijda's solution to the issue of the natural and appropriate units and levels

[77] To this he adds that the general construal also remains relatively constant regardless of the specific manners in which the situations are re-appraised over time and that this relative constancy is plausibly what accounts for the categorisation of the entire episodes under single specific emotional categories. As I have suggested in chapter 1 there may be several factors underlying the classification of multi-attitude emotion episodes under global emotional categories and my position on this question will be 'pluralist'.

of analysis of emotion is thus that there are various equally adequate candidate emotional particulars on which we may focus our attention provided that we do not lose sight of their being composed of, or embedded in, one another.

> Successive emotions within an episode share some components and vary in others. It still makes sense to say "I first was amazed, then it turned into anger, then into utter confusion," because, at that level of analysis, the terms refer to the temporary predominance of a given appraisal and / or a given action readiness. "An emotion", in this sense, may last as long as a given mode of appraisal or a given mode of action readiness is maintained; but how long that is can only be indicated by approximation because of the blendings-over and thinkings-back. At the same time, the nature and composition of each emotion is shaped by the […] context in which it occurs. Each individual emotion occurrence carries the stamp of that context, just as every pleasure does. (op. cit., 195).

Now, Frijda is not especially interested the question of emotion-individuation as a question of ontology since he takes it to be ultimately terminological. His view is thus not that specific emotion concepts should be reserved to emotional particulars identified in reference to particular appraisals, or action tendencies; it is "sheer desire for order to reserve the word 'emotion' for an episode concerned with a particular relational theme, or for a phase dominated by a given form of action readiness" (ibid.). It is, however, fairly clear that Frijda would take the concepts of 'emotion episodes' and 'emotion attitude' to be distinct and not necessarily *coextensive*. This does *not* mean that emotion episodes never coincide with episodes of a single kind of emotion: we do often undergo emotion episodes of which we have an immediate categorical awareness as, episodes of 'fear', or 'anger', or 'shame' etc. The fact that the two concepts are not coextensive, however, is sufficient to motivate the inclusion of multi-attitude emotion episodes among our targets and try to provide an analysis of such phenomena in their diachronic complexity.

Although I haven't shown conclusively that multi-attitude emotion episodes are emotional 'wholes' or 'unities' as opposed to being sequences of specific episodic emotions, I take this chapter to have cast serious doubts on the rationale for the tendency to individuate our emotional particulars in reference to the obtaining of single specific emotion attitudes and to have gone some way towards defending the plausibility of this approach by identifying precise criteria for the individuation of emotion episodes. This should suffice to motivate an attempt to consider emotion episodes as emotional particulars and raise central questions of ontology independently of the assumption that emotional experience is to be particularised in reference to the retention on the part of the emoting subjects of specific emotion attitudes.

As we turn to multi-attitude emotion episodes as our targets, various issues come into view which would not come into view if we were limiting our concern to emotional particulars individuated *via* single specific emotional categories. The first question to be addressed in view of my aims is the question of 'emotional persistence': if we accept that emotion episodes may constitute emotional continuities independently of their being instances of some single emotional category, we face in fact the issue of explaining the supposed continuity underlying the variety of phases their complete lifecycle involves. The idea here is that in the simpler instances, that is, when emotion episodes happen to coincide with single specific emotions, the question of emotional persistence is answered straightforwardly by appealing to conditions of attitude-*retention*: for example, an episode of fear persists or continues if the relevant attitude–i.e., fearing–is retained, that is, if the emoting subject continues to fear the relevant situation. But in cases in which different emotion attitudes succeed one another, if we accept that such cases are cases in which single emotion episodes span changes in attitude, we are pressed to ask what else could account for the continuity of such episodes if not the retention of one specific attitude. In the next chapter, I will raise the question of emotional persistence and, by answering it, I will begin to develop my analysis of the causal structure of emotion episodes as composite processes over time–i.e., the structure of causal relations

underlying their occurrence and persistence. The main upshot of my analysis will be that such causal structure, unlike the causal structure of other kinds of affective reaction, crucially involves our sustained attentional engagement with emotional situations. To clarify further, the scope of the analysis will not be limited to multi-attitude emotion episodes. Some emotion episodes are very short, and they persist without much variation in quality, intensity, or complexity, as when noises cause bouts of fear which peak quickly and quickly subside. Other emotion episodes are caused by more protracted *engagement* with emotional situations, and they persist while they vary along several dimensions. If we accept that at least some emotion episodes are inherently articulate in this way, that is, that they cannot be analysed into *mere* sequences of specific episodic emotions, we are forced to look for an account of emotional persistence which may apply to all episodes, not just to the simplest instances.

Chapter 3

THE CAUSAL STRUCTURE OF EMOTION EPISODES

Introduction

A very natural starting point in view of addressing questions of causation about emotions, and one which has the advantage of being shared by the vast majority of theories of emotion independently of how their chosen emotional particulars are individuated, is the classification of the latter as kinds of (affective) 'reaction'.[78] The concept of 'reaction' is a *specific* causal concept and, as we will see, it puts specific constraints on an analysis of the causal structure of the phenomena to which it is taken to apply, especially as concerns the way in which they persist or unfold over time. Now, the first and most obvious implication of the concept of 'reaction' is an implication of passivity: reactions are elicited or caused in us from *without*, viz., *passively*, as opposed to being brought about by us. Accordingly, the reactional character of emotions is commonly mentioned to emphasise their passive character. Tappolet (2017), for example, uses the idea that emotions are reactions to support the analogy she wishes to

[78] As mentioned in the introduction, some authors have criticised this view and have offered radical alternatives – such as views in terms of actions or strategies (Sartre 1939; Solomon 1976) or 'skilled cultural performances' (cf. De Sousa 1990). I shall leave aside this sort of approach (see Introduction, Significance). Strictly speaking, *perceptual* theories of emotions would also lie outside of the scope of the following discussion: as we will see, the causal structure of perceptual experience is crucially different from the one entailed by the concept of a 'reaction.' However, the relevant concept of 'perception' is often broadly construed in such theories, in ways that do are not incompatible with the view that emotions are reactions (whether this is an issue for the analogy between emotions and perceptual experience on which the view rests is another question). To wit, some of the main proponents of such an account of emotion such as De Sousa (1987), Prinz (2004), Tappolet (2000; 2016) either explicitly qualify emotions as reactions, or, at any rate, seem to use 'emotions' and 'emotional reactions' interchangeably.

defend between emotional and perceptual experience, and more specifically, in suggesting that this analogy holds with respect to the passivity of the two kinds of experience:

> [Y]ou can neither decide to feel fear when you do not happen to experience this emotion nor chose to see white snow as orange, say. Though there are indirect ways to control our emotions, such as breathing slowly to avoid panic, or going for a walk to attenuate one's anger, emotions are *passive* states. […] In general, emotions, whatever their kind, automatically arise in response to the world. *They are reactions*, *as opposed to things we do*." (20, italics added).

I will suggest instead that while classifying emotions as affective reactions does imply that they are elicited (or evoked or triggered or kicked-off), that is, *initiated* by factors lying outside the scope of direct agential control, is does not imply they are wholly passive phenomena. Reactions, unlike other passive occurrences, such as, e.g., perceptual experiences, have what I will call a 'proper course' over time: having been elicited, they may then persist or unfold over time in ways that are partially independent of their initiatory elicitation, that is, the way they then *run their course* may depend on other causal factors. Notice how this idea connects with the idea that emotion episodes are dynamic in that they have lifecycles, or natural temporal trajectories characterised by certain stages or phases of development. My discussion of 'proper course', and more specifically, the proper course of emotion episodes as kinds of affective reaction, will aim to clarify what underpins such lifecycles from the point of view of their causation. In principle, this partial independence of reactions from their initiatory elicitation does not itself make them any less passive. In fact, other passive factors might happen to explain the proper course of given reactions–think, for example of some kind of inertial progression towards subsidence, such as the gradual return to baseline of the difference between one's core temperature and skin temperature which takes place when one enters a warm house after having been exposed to cold temperatures. It does, however, make

some room for agential factors to enter the scene, and determine the way in which the reaction persists or unfolds over time–as they do, for example, in bodily action we tend to treat as reactional, e.g., dodging a blow to the head. In other words, the evident implication of passivity of the concept of 'reaction' is limited to the distal initiatory elicitation of the phenomena to which it is taken to apply. This fact that has been widely overlooked, and the issue with this neglect is that the view that emotional reactions as *elicited* events are passive, though correct, captures only partially the relation between emoting subjects and the emotional situations, and thus needs to be supplemented if we are correctly to theorise our capacity for emotion: emoting subjects, not only react affectively to emotional situations as distal elicitors, but they also *engage* with emotional situations as *objects* of their current interest and sustained attention, and this engagement plays a role in the emotions' continued causation. A key aim of this chapter is thus to establish that classifying emotional particulars as kinds of (affective) reactions does not imply that they are wholly passive phenomena. This will prepare the ground for the view that emotion episodes are both episodes of affective reaction and episodes of sustained emotional engagement. Towards the end of the chapter, I will defend this dual view of emotion episodes by contrasting their causal structure with the one instantiated by other kinds of affective reaction.

It will be useful to begin by locating my analysis of the causal structure of emotion episodes in relation to debates on emotion causation. The classification of emotional particulars as kinds of (affective) reaction is often considered uncontroversial;[79] indeed, it is common to introduce the idea that emotional particulars are reactions as part of some preliminary characterisation of such phenomena as opposed to a thesis to be developed and defended.[80] Starting with the idea that emotional particulars are essentially reactions, what then remains

[79] The term 'emotion', when referring to *occurrent* emotion, as opposed to single-track emotional dispositions or sentiments, and the phrases 'emotional reaction' or 'emotional response' are usually employed interchangeably, and so are, for related reasons, the phrases 'emotion causation' and 'emotion elicitation.'

[80]Müller (2017) Mulligan (2007) and Zamuner (2015) are the only authors, to my knowledge, to offer sustained argumentation in defence of the view that emotions are reactions or responses.

to be done is to provide an account of emotion causation distinguishing emotional reactions from other kinds of affective reactions and differentiating the various specific emotions. Both these issues are often taken to be solved by specifying what emotional reactions are reactions *to*, that is, by identifying their *distinctive* elicitors or triggers.[81] According to all broadly cognitive accounts to emotions, emotional particulars differ from other affective reactions such as pain or sensory pleasure in that, while the latter are immediately caused by certain kinds of sensory stimulation, the former are reactions to cognition–and more precisely, *cognised* emotional situations. This follows from the claim that emotions are not independent 'modes of access' to the variety of worldly objects they seem to be about (cf. Teroni 2017), so that for one to react emotionally to such objects, the latter must somehow be cognised–i.e., noticed, imagined, remembered–*prior* to the occurrence of the emotional reaction.[82] While they share this core presupposition, cognitive accounts diverge as regards the precise nature of the emotion-eliciting cognition, and most notably, they diverge on whether it involves axiological information–e.g., whether the fearsome or shameful character of the situation is appraised by the subject *prior* to her reacting to the situation with fear or shame.[83] This is an issue on which I remain neutral. Secondly, since many emotion theorists are interested in how to differentiate the various specific emotions, much effort has been put into identifying the particular cognitive contents which elicit them, so as to differentiate, say, 'fear' from 'anger' or, to take a more subtle distinction, 'regret' from 'sadness' or 'nostalgia'. This is also an issue to which I shall have little to contribute.

While the issue of the precise nature of the emotion-eliciting cognition and the issue of which particular cognitive contents elicit which emotions are both much debated, the implications

[81] Cf. Moors (2009) very useful review of the way in which different theories of emotion address what she calls the '*elicitation* problem'.
[82] That emotions are cognition-mediated in this sense is the central premise of all so-called 'appraisal theories' (Ellsworth & Scherer, eds., 2003). For compelling discussions of its advantages over non-cognitive alternatives, see Reisenzein and Shönpflug (1992) and Teroni (2017).
[83] Cf. footnotes 45 and 47.

of the very classification of emotional particulars as (affective) reactions are rarely discussed. One of my aims is to fill this gap. As I have anticipated, I will argue that this classification implies a partial independence of the emotions' continued causation over time from their initiatory elicitation. If this is correct, questions of causation about emotion cannot be taken to amount to questions regarding the precise nature of their elicitors as normally assumed.[84] I will argue that emotion episodes differ from other kinds of affective reaction not only in their specific elicitors, but also in that their causal structure is more *complex*, or more precisely, in that there is a *deeper* causal structure we need to uncover if we are to correctly theorise emotional persistence as distinct from other kinds of persistence of affect. This further stratum of complexity is introduced by the fact that emotional reactions are reactions to cognised situations, and their being *cognitively* mediated 'complicates' their causal structure. Firstly, the causal structure of emotions, though similar to that of other affective reactions in that it involves an initiatory elicitation of affective changes, differs from it in that the course of emotional reactions inherits certain structural characteristics from the way in which the various agential cognitive processes they involve unfold over time. Secondly, the relation between affect and cognition is–unlike the one between affect and sensation– a relation of reciprocal causation. There is much here that needs to be unpacked, but this should suffice to outline the scope of what follows and clarify how it relates to current debates on emotion causation. I will now introduce some preliminary assumptions, to be defended in due course, concerning the nature and structure of affective reactions. This will allow me to introduce the distinctions and terminology I will rely on.

[84] This line of reasoning mirrors closely the one sketched by Steward (2013) in defence of the significance of the concept of an *individual* process: "But what is the cause of a token process, conceived of as an entity with a robust form, a normal course of development, something which can obtain extra temporal parts without detriment to its continued identity? Not only what triggers it, but also what sustains it, what keeps it on course, what prevents it from ceasing or disintegrating. […] For here we have to explain not only the occurrence of its front end, as it were, but also its course, and for that we need a different sort of explanation. To explain its initiation is not to explain everything that needs to be explained." (810). The way in which I raise the question of persistence in §3.2 owes much to Steward's suggestions.

1 The classification of emotional particulars as kinds of affective reaction

The category 'affective reaction' is an overarching affective category comprising many kinds of affective episode alongside emotion episodes, such as, e.g., reflexive pain reactions, sensory pleasure and discomfort, pleasure in ongoing physical and cognitive activities, phenomena such as thirst, hunger, and other physical cravings. These disparate phenomena qualify as 'affective' because they involve affective changes, that is, positive or negative hedonic feelings, and, relatedly, given motivational states.[85] For example, some kinds of pleasure involve a motivation to preserve one's engagement in a pleasurable ongoing activity, pain often involves a motivation to protect the body-part hurting and hunger involves a motivation to rid oneself of certain unpleasant sensations by means of eating. As for their classification as 'reactions', it is natural to assume that it is due to something pertaining to their causation–at first approximation, it is due to their being caused from without by the presence of their elicitors. As we will see, all affective reactions are complex episodes, for they all include both given affective changes and the sensations or cognitions from which they *inherit* their distinctive intentionality.[86] For example, pain episodes involve located sensations from which they inherit a body-directed intentionality whereas emotion episodes involve occurrent cognition from which they inherit their self-directed or world-directed intentionality.[87] Employing a widespread terminology which originates from Brentano's ontology of the mind,[88] I will call the sensations or cognitions involved in affective reactions their respective *bases*. Within the sub-class of affective reactions that are based on sensations,

[85] Following Tomkins (1994), I am assuming that hedonic feelings are intrinsically motivating in virtue of their inherently rewarding or punishing character. Teroni (2018) for discussion of the nature of affective valence.
[86] Deonna & Teroni, (2012); Meinong (1917/1971, 26-27, 60-61); Mulligan (2007); Reisenzein & Schönpflug (1992) and Teroni (2007) for discussion of the claim that affective reactions *inherit* their intentionality from presentations, i.e., sensations and cognitions.
[87] Teroni (2007), for a discussion of the intricacies of the relation between cognitive presuppositions and emotion and the implications of different accounts of this relation for the issue of emotional intentionality.
[88] See Mulligan (2004): "In his *Psychology*, Brentano notes that language suggests that certain emotions relate to objects–we say we are sad or upset about this or that. In such cases emotions "relate to what is presented in" the presentation they are based on (PES-E, p. 90, PES-G, p. 126). In other words, the intentionality of emotions is *inherited* from that of their *bases*, presentations and, in some cases, judgings." (81, italics added).

we may further distinguish cases in which the sensations are caused 'exogenously', that is, by an external stimulus, and cases in which they are caused instead 'endogenously', that is, by an internal stimulus generated by workings of physiological mechanisms that preserve our homeostasis or physiological equilibrium.[89] We may thus distinguish three types of affective reaction in terms of their bases: 'sense-based' affective reactions, such as pain and sensory pleasure; 'drive-based' affective reactions, such as thirst, hunger and other physical cravings, and cognition-mediated affective reactions, or emotional reactions. One last assumption regards the nature of the 'basing' relation. I will take this sort of relation to be partly causal: the affective changes involved in all affective reactions are causally dependent on their bases. As we will see, the basing relation is not, however, merely causal. It is also more 'substantial' (Teroni 2017) or constitutive: unlike other kinds of effect or result, the affective components of affective reactions cannot outlast their bases in the way they may outlast their distal elicitors, and they are in part to be specified in terms of them (cf. Soteriou 2013, 70-75). To illustrate, the affective changes involved in an episode of pain–that is, what we may call the subject's *suffering* or her state of *being* in pain–are not only causally, but also constitutively dependent on bodily located sensations: the suffering, if it is elicited, say, by entering in contact with a hot stove, is in part to be specified by appealing to the relevant sensations located in the part of the body that was in contact with the stove, that is in ways that make reference to their spatial and temporal features and their distinctively 'painful' quality–e.g., 'cutting', 'drilling', 'burning' etc. The same goes for the basing relation involved in emotion. As we have seen in the previous chapter, emotional particulars, unlike specific mood changes, do not outlast the termination of the cognisance on which they are based. It is worth emphasising what I mean by 'cognition' is occurrent cognition–viz., ways of attending or ways of having one's attention occupied. To say that emotions have cognitive bases is not to say that they *presuppose* given cognitive states, such as, for example, a belief that if I miss my

[89] See Cannon (1932).

deadline I will be in big trouble. This is the basis of my *standing* fear of missing my deadline, but–fortunately–I only feel my fear of missing my deadline intermittently, i.e., any time that the thought that I may miss my deadline happens to cross my mind. This point is nicely illustrated by Stumpf (1899) in the quote below.

> I would not admit […] that the intellectual states are only causes of emotions. […] Envy comprises the presentation (Vorstellung) and judgment of the good concerned as one belonging to the fellow human, and *it exists only as long as these intellectual elements do exist; they belong to its substance. If the presentation disappears or the judgment changes, the affect disappears or changes as well*, even though in many cases aftereffects and sensory repercussions may remain. (58, in Reisenzein & Shönpflug 1992, 36, italics added).

Much of what I have said will be clarified and fully defended over the course of the chapter. With these preliminary assumptions and terminology in place, I may now turn to the key question of this chapter, namely the question of emotional persistence.

2 The question of emotional persistence

What causes emotion episodes to persist and unfold as they do? What is the causal process or mode of production underlying their continued causation? This is the issue that will occupy me in what follows.

Questions of causation focussing specifically on the persistence of the emotional particulars are not often raised. Since emotional particulars are commonly recognised to have duration or temporal extension, it is natural to wonder whether there is some reason for this neglect. We may start by noticing that this neglect seems to go hand in hand with a given framing of questions of causation about emotions, a framing that equates the general concept of 'causation' and the concept of 'elicitation'. In order to appreciate the link between the neglect

of the phenomenon of emotional persistence and the tendency to assimilate the causation and the elicitation of emotional particulars it should suffice to notice that specifying elicitation conditions for occurrents with temporal extension is tantamount to specifying elicitation conditions for whatever punctual occurrence is taken to 'mark' their onset at an instant: questions of elicitation are answered by specifying what triggers or kicks off our emotional reactions. Consequently, framing the problem of emotion causation in terms of elicitation suggests that for all that it matters to an account of emotion causation, emotional particulars may as well be devoid of duration. While this does not imply that emotional particulars are punctual, duration-less, phenomena, it does raise the question as to whether and how an account of emotion elicitation, that is, an account whose scope seems to be limited to the causing of the onset of emotional particulars, might yield an account of emotional persistence. In next sub-sections, I will address this question by offering an analysis of the concept of an affective 'reaction.' My suggestion will be that what warrants, supposedly, the tendency to bracket or abstract from the temporal diachronic dimension of emotional particulars is the particular interpretation of the concept of 'reaction' that I will call the Passivist Reaction Ontology (PRO), an interpretation on which an affective reaction's duration does not pose any problems additional to the ones posed by its initiatory elicitation. In the next section, I will provide an account of the concept of 'reaction', and I will begin to argue that it does not exclude agential factors from being at play in the causation of the phenomena to which it is taken to apply.

3 The concept of reaction

What do we mean when we qualify certain occurrences as reactions? The general concept of a 'reaction' is a *causal* concept: part of what we mean when we speak of an occurrence as a reaction is that such an occurrence is an effect or result of the causal impact of something,

namely the elicitor of the reaction, on something else, namely the subject of the reaction. This can be made more precise as follows. 'S reacts to E' admits of two different readings: an 'occurrent' one and a 'dispositional' one. To say that S reacts to E on the occurrent reading is to say that a reaction is occurring, whereas to say that S reacts to E on the dispositional reading is to say that S possesses a disposition to react in a certain manner, when acted on or affected by E–in other words, it is to say that S is 'sensitive' to E. Now, such sensitivities come in degrees, and whether S will react to E, as well as further graded or dimensional properties of the reaction in question, such as its extent or intensity, shall partly depend, plausibly, on the degree of S's sensitivity to E. If the degree of sensitivity possessed by S is known, then, in order to know whether S will react to E, we just need to know whether E is strong enough an elicitor to reach and exceed the subject's reactivity threshold; we just need to know whether the elicitor has the strength required to engage the subject's sensitivity and allow the latter to become manifest. The manifestation of the subject's sensitivity we observe if the subject's reactivity threshold is reached *is* the subject's reaction.

When we speak of an occurrence as a reaction, we are thus qualifying such an occurrence in reference to the place that it occupies in this sort of structure of causation, and such place seems to be that of an *effect* or *result*. Now, does this mean that 'reaction' and 'effect', or 'result' have the same exact meaning, or is there a sense in which reactions are *special* kinds of effect? In the next few sections, I will argue that reactions are special kinds of effect, and I will try to identify the relevant difference. I will then apply my account to affective reactions.

I will start by drawing on an analysis of the concept of 'reaction' recently offered by Zamuner (2015) who explicitly endorses the view that reactions differ from other effects. Zamuner attributes to reactions five main features: polarity, passivity, duration, time-course, intensity. Although he does not take these features to be 'defining'–that is, necessary and sufficient – he seeks to establish that emotions are reactions by showing that they instantiate all of them.

The features on which Zamuner mostly insists are, firstly, 'polarity' and, secondly, 'passivity.' While I agree that emotions are reactions and that they have the above features, I do not think that they are reactions in virtue of these features. In particular, they are not reactions in virtue of having *polarity* or valence. Moreover, 'passivity,' 'duration' and 'time-course' are instantiated by emotional reactions in ways that are importantly different from the way they happen to be instantiated by other, non-affective reactions Zamuner utilises as models–i.e., reactions such as, e.g., allergies and infections as well as tickles and other contact sensations. Let me begin by noting, following Zamuner, that 'reaction' applies to many kinds of occurrence, including bodily action, aspects of the mind and physical conditions.

> Sometimes we describe *behaviours* and *actions* as reactions. Spontaneously withdrawing your hand from a hot stove is a reaction, and so is arching your shoulders when a drop of water falls down your neck. But equally, we describe *mental states* as reactions. Fear is a reaction to a danger; itching a reaction to an irritant. And of course, we can also describe *physical conditions* as reactions. Hay fever is a reaction to pollen and tanning a reaction to sunlight. (Zamuner 2015, 24, italics added).

The question we thus need to address is what these disparate occurrences share that might distinguish them from other effects. Following Mulligan (2007), Zamuner (2015) proposes that talk of reactions differs from talk of simple effects in that it implies what he calls *polarity*:

> As Mulligan (2007, p. 26) notes: "having a valence suffices to make a state a reaction. States of opposed valence or 'sign' constitute opposed reactions." In my terminology, reactions have *polarity*. For example, we can think of assertion and negation as polarly opposed linguistic reactions just as we can think of worsening and improvement as polarly opposed physiological reactions to a pathogen or drug treatment. By contrast, effects need not have polarity. We can speak of x being the effect of y without implying that x has any polarity. Of course, we can say that the treatment had a positive effect on Jones'

> condition. But equally, we can say that Jones reacted well to the treatment. The difference is that while talk of reactions typically implies polarity, talk of effects tends to be neutral. (ibid., 24).

The core idea here is that when we call effects 'reactions', we advert to the fact that some polarly opposed reaction could have taken place given the very same elicitor. For example, one may react to a given piece of evidence with either belief or disbelief (cf. Mulligan 2008). A first problem with this idea, that Zamuner, following Mulligan (2008, §5), acknowledges, is that not all reactions have opposites. Grief, guilt, pride, shame, enjoyment, and other affective reactions, for example, lack (nameable) opposites. Zamuner's solution is to construe the polarity of affective reactions in terms of the presence of tendencies to approaching *vs* avoidant behaviour. To forestall immediate objections, he clarifies that this distinction is "independent from, and neutral to, the distinction between positive and negative emotions" (28). There are many other ways to spell out the polarity or valence of affect.[90] I will argue that regardless of the way in which it is spelled out, this sort of strategy, if it is meant to yield an account of the concept of reaction, it problematic, for the very fact that polarity might be spelled out in a number of manners and that these manners do not always produce results that may be aligned in the right way, casts doubt on the view that polarity is a feature that affective reactions possess in virtue of their being reactions. In the next section, I will discuss 'polarity' as a criterion for reactions and I will suggest that emotional reactions have polarity not in virtue of their being reactions but in virtue of their being affective phenomena. I will then suggest that emotion episodes are reactions in virtue of having a 'proper course', that is, the property of *outlasting* their distal elicitors, or more precisely a certain independence from their initiatory elicitation. This will later allow me to prepare the ground to reframe the relation between the concept of a 'reaction' and the categories of the 'agential' and 'passive.'

[90] For a review of the available options in theorising affective polarity see Colombetti (2005). Teroni (2018) also reviews the options and argues for the view that polarity or valence is essentially connected to evaluation.

4 The valence or polarity of (affective) reactions

In this section, I will raise some issues with the idea that polarity is a criterion for reactions. First of all, although it does make sense to speak of affective reactions as positive or negative, it is not clear that all effects we classify as reactions have polarity, and, hence, it is not obvious that affective reactions have polarity in virtue of their being reactions. Chemical reactions–e.g., combustion–do not have polar opposites, and the same goes for *physiological* reactions such as allergies or infections. Zamuner addresses this problem by suggesting that we may think of "worsening and improvement as polarly opposed physiological reactions to a pathogen or drug treatment" (op. cit., 24). But the terms 'worsening' and 'improving' do *not* pick out *kinds* of reaction; rather, they refer to ways in which one may appraise given effects. Any effect, given some framework of evaluatively connotated expectations establishing which kind of effect is desirable and which kind of effect is undesirable, might be said to be a way for the relevant situation to worsen or improve, so that in introducing these concepts Zamuner somewhat trivialises the notion of polarity. It is hard to see how polarity or valence might be taken to be necessary for reactions–as Zamuner concedes. Is polarity or valence sufficient? Perhaps polarity is a feature of *psychological* reactions that is not and should not be expected to be instanced by chemical or physical reactions. If this is the case, as Mulligan (2007) suggests, perhaps having polarity suffices to qualify as a reaction in the mental domain. In defence of this claim we may observe that many psychological occurrences we classify as reactions do come in polar opposites–e.g., belief and disbelief, many occurrent emotions, hunger and satiation, pain and pleasure etc.–and reactions for which we do not have exact nameable opposites may be accommodated by suggesting that they still stand in contrast with some 'relevantly neutral condition': ascription of affective reactions suffices to exclude what we may call affective *indifference*, that is, the obtaining of the relevantly neutral condition in the presence of elicitors which might plausibly be expected to produce *some* affective reaction–consider, for example, "the loss left her indifferent." This point is connected to

Mulligan's (2007) observation that although one would typically react to seeing something or seeing that p with either belief or disbelief, the view that such attitudes are reactions also implies that it is always a *contingent* matter whether one will react doxastically in the presence of the relevant elicitor: "[o]ften, but not always, we respond doxastically to seeing something or seeing that p" (219). However, it is not clear that in all such cases the notion applies in the same exact sense. Indeed, it is not clear what is meant exactly by 'polarity' until we can spell it out, for example, in terms of being positive or negative for our organismic or psychological functioning, involving a tendency to approach or avoidance behaviours, including positive or negative hedonic feelings. As soon as we spell out polarity in any one of these manners, we see that these possible meanings of the term 'polarity' are not always aligned in the way we would expect. For example, purely somatic reactions like shivering, or coughing, or vomiting are positive reactions for the organism–the organism, in undergoing them, is reacting desirably to certain elicitors. On the other hand, they involve avoidant behaviour, and they are themselves unpleasant so that what is 'good' for the organism is *not* 'good' hedonically. Similarly, drive-based reactions, like thirst, hunger, and other cravings, are good for the organism in that they serve essential organismic functions, and they do this by motivating seeking behaviours, i.e., kinds of approaching behaviours. On the other hand, arguably, they do so in virtue of their negative hedonic polarity.[91] One final consideration which should lead us to rethink the nature of the relation between the concepts of 'reaction' and 'polarity,' is that it is not clear that affective reactions possess polarity *in virtue of* their being reactions as opposed to their being *affective*. Sentiments, such as love and hate, also come in polar opposites, and so do some traits of character and personality, such as courage and cowardice, or introversion and extraversion. Nonetheless, we do not classify standing dispositions as kinds of affective reaction: while they dispose us to react in given manners,

[91] The felt *attraction* towards what is sought–e.g., appetite or more generally *appetitive* desiring–is perhaps positive, but it is best viewed, arguably, as a distinct affective reaction that drive-based reactions, so to say, 'recruit' to further their function–cf. e.g., the way in which pain *recruits* fear (see Ombrato & Phillips 2020).

they themselves do not seem to have the sort of causal structure entailed by the concept of 'reaction.' First of all, they do not seem to have the right ontological properties for they are kinds of state and not episodes; secondly, and relatedly, they are not caused by an elicitor at an instant, but they come to obtain in many ways, and, typically, they gradually establish themselves over time (cf. Wollheim 1999, ch.1). Thirdly, although the *acquisition* of standing dispositions–e.g., my attraction for Ambra–may sometimes presuppose possession of other dispositions that got established in the past–e.g., my attraction for red haired women–not all of our standing dispositions manifest other standing dispositions located further 'upstream', so to say, but as we have seen, reactions essentially manifest dispositions or sensitivities. Finally, although the inception of our standing dispositions, as Wollheim (1999, 2) highlights, is caused by–and *marked* by–given occurrences–e.g., "waking up and seeing a frog standing on his chest could establish in a boy a lasting terror of frogs." (2)–intuitively, this is an instance of acquisition and not elicitation–i.e., having been *initiated*, dispositions do not need to be sustained in any way to continue.

To conclude, while affective reactions do typically possess polarity or valence, it is not clear that this is essentially connected to their being reactions. Indeed, it seems rather due to the fact they are *affective* phenomena. This gives us a reason to look for some other criterion for affective reactions. I will now suggest that the most attractive candidate is their distinctive temporal character.

5 The distinctive temporality of (affective) reactions

As I have mentioned, Zamuner grounds the claim that emotions are reactions in three other features besides polarity, namely passivity, intensity, duration, and time-course. It will be useful to begin by presenting and discussing the way in which he understands these features and the relations they bear to one another because it is emblematic of what I will call the

Passivist Reactions Ontology (PRO). Zamuner does not offer a definition of the relevant concept of 'passivity' or 'involuntariness.' However, he makes it clear that he understands it on the model of the passivity of physical reactions such as allergies, infections and contact sensations. To qualify his claim he specifies that emotional reactions are only *typically* involuntary, since he takes it that at least some emotions, and most notably, 'gratitude,' allow for direct agential control.[92] He also observes that the fact that emotions are involuntary is not incompatible with the possibility of exercising indirect agential control over them, as we may do, for example, "when we try to *calm* our anxieties, when we try to *prevent* ourselves from falling in love, when we *allow* ourselves to give in to sadness, or when we *let* ourselves get carried away with anger" (26, italics added). The other properties, and the way in which they interact, are elucidated in the passage I quote below.

> There are other similarities between emotions and non-emotional reactions. First, they both have *intensity*, *duration*, and *time course* (understood as the way in which a reaction unfolds over time). Allergies typically arise as soon as exposure to an allergen occurs and subside as the allergen is removed and the levels of inflammatory mediators return to baseline. Similarly, most emotions have fast onsets whereas their offsets can vary from fast to slow. Surprise, for example, arises and subsides quite quickly whereas anger can arise quickly but may take time to subside. Regret, on the other hand, builds slowly and lasts longer. Evidence indicates that the intensity of a reaction at onset can determine its duration and time course. While this is particularly true of physical reactions (e.g. Spickett 2006), there is also converging evidence from studies on emotions (e.g. Verduyn *et al.* 2009). (op.cit., 30).

[92] "This seems particularly true of gratitude, for we may decide that our enemies deserve gratitude, just as we may decide that our friends deserve none. As Strawson (1974, p. 6) aptly notes, gratitude is an emotion that we may decide to grant or withdraw. In line with this view, one could imagine situations in which gratitude is not so much a reaction, but rather a response to another person's actions." (op. cit., 26).

While Zamuner is right to point out that emotions and non-emotional reactions both have intensity, duration, and time-course, it is not clear that intensity, duration, and time-course of emotional reactions are to be understood on the model of allergic reactions, tickles, pains, or any other physical reaction and contact sensation. Zamuner says that allergic reactions "arise as soon as exposure to an allergen occurs and subside as the allergen is removed and the levels of inflammatory mediators return to baseline" (ibid.) and suggests that emotional reactions are similar in this respect. In doing so, he not only suggests that emotions have duration, that is, an onset and an offset, and given time-course, that is, ways in which they unfold over time from their initiatory elicitation to their complete subsidence, but also that the duration and time-course of both allergic reactions and emotional reactions are underpinned by passive factors, namely initiatory elicitation of certain changes and inertial development towards subsidence–that is, a gradual return of the relevant changes to baseline. In this, Zamuner's proposal is an instance of PRO: our emotional reactions are not only caused, but also causally sustained throughout their duration by factors lying beyond agential control, and they are, hence, entirely passive.[93] That Zamuner's understanding of what I have called emotional persistence is based on PRO is confirmed by the fact that he explicitly connects the duration of emotional reactions to their *initiatory* elicitation–"evidence indicates that the intensity of a reaction at onset can determine its duration." In what follows, I will try to undermine this connection.

Let me begin by trying to clarify what is meant by *intensity* in the case of emotional reactions. Frijda (2007, ch.6) discusses empirical evidence showing that emotional intensity or strength is not to be thought of as "something elementary like the loudness of sounds" (155). Indeed,

[93] We can also see from the quote that Zamuner understands duration and time-course of emotional reactions in a way that I have already problematised in §2.3 in my discussion of Pear's view of the duration of emotions, namely as a *species-specific* psychological fact: different emotions happen to possess different durations and time-courses. I shall thus ignore this element of Zamuner's proposal and focus on the view he puts forward as regards temporal features shared by all emotional reaction, that is, their duration and time-course understood as progression towards subsidence.

the overall reported intensity of emotional experience as rated by the subjects of his studies turns out to be a summary judgment referring to various factors, among which, notably, (i) strength of somatic feelings and felt arousal, (ii) strength of the relevant action-tendency, (iii) extent of the long-term impact of the emotional experience in terms of changes in one's beliefs and subsequent behaviour, and (iiii) recurrence in thought of the eliciting stimulus. Now, in Fijda's study, unlike the ones on which Zamuner relies–i.e., Verduyn et al (2009)–duration does not turn out to correlate with any of these factors, nor with the overall intensity of emotional experience: "[d]uration, as rated by the participants, correlates weakly or not at all with the other factors, including overall felt intensity. Estimated duration of emotions thus appears to be nearly independent of felt intensity[94]" (Frijda 2007, 156). It is important to notice that in the study discussed by Frijda neither overall intensity nor any of its various dimensions were rated at onset: the subjects' reports concern intensity of the whole emotional experience from beginning to end. Therefore, they do not conflict with evidence such as that on which Zamuner relies, which concerns, instead, the *initial* intensity. Frijda's evidence, however, suggests that other factors must be at play, and mediate between intensity at onset and the duration of the emotional reactions: if no correlation is observed as soon as we move from considering intensity at onset to considering overall intensity, the duration of the reaction cannot simply be a matter of gradual subsidence or return to baseline on the model of non-emotional reactions. What might such factors be? Verduyn *at al.* (2009), Zamuner's source of empirical support, discuss one option. Pace Zamuner, in fact, they do not take intensity at onset to *directly* explain duration. The main explanation they hypothesise of the fact that intensity at onset correlates with duration is the following: the *importance* of the eliciting situation–viz., the importance of the underlying concerns and the magnitude of the relevance of the eliciting situation to such concerns–influences *both* the intensity at onset

[94] Fredrickson & Kahneman (1993) obtain the same results with an on-line study and present further evidence from animal studies.

and the frequency of what they call mental 'reappearances' of the eliciting stimulus–namely, how often the subject reports thinking *back* to the eliciting stimulus (87-90), and this is so for the latter "reactivate" the emotion. In further studies (2011; 2012; 2015), Verduyn *et al.* confirm this hypothesis and refine it by adding that for such mental reappearances to prolong the duration of emotion episodes, the eliciting stimulus must mentally reappear in ways that reactivate the original appraisal, as opposed to reappearing in reappraisals such as the ones involved in emotion down-regulation. The fact the that intensity of the reactions at onset only *indirectly* influences their duration, that is, *via* mental reappearances of eliciting stimuli, and, hence, the renewal of the original appraisal, makes room, in principle, for agential factors to enter the scene. This is how Verduyn *et al.* (2011) actually interpret their own findings: "people generally do not passively undergo their emotions; rather they often actively influence their emotional experiences (20-21).

Verduyn's conclusion is a bit quick, however, for not all cognition is active and it is not clear that such mental reappearances are caused by agential cognitive processes as opposed to *unbidden* thoughts or mere ideational associations. If mental reappearances turned out to be caused by passive factors, same in kind to those which cause the reactions' initiatory elicitation, PRO would be ultimately vindicated. A model for this sort of cognitive causation is provided by *trauma effects*: the more intense the initial emotional reaction, the more sensitive the subject becomes to the eliciting situation, the more likely it is that her initial reaction be elicited anew–until something like fatigue, distraction, habituation, or active down-regulation intervenes. Frijda (2007) captures this phenomenon, of which trauma effects are extreme variants, in one of his laws of emotion–the Law of Conservation of Emotional Momentum: "*emotional events retain their power to elicit emotions indefinitely, unless counteracted by repetitive exposures that permit extinction or habituation, to the extent that these latter are possible.*" (14). The idea is that once we react to given situations by undergoing emotional experience, we become indefinitely susceptible to what Verduyn calls their mental reappearance. In other words, we

acquire a standing emotional sensitivity which tends to cause, for example, certain mental images, thoughts, or instances of noticing. In discussing his law, Frijda says that it "will be difficult to prove, because it asserts resistance against change when nothing happens […]. Yet there is ample clinical and anecdotal evidence to support it.

> "The persistence or recurrence of trauma effects is […] well known. Emotions surge up when stimuli resembling the original stimuli are encountered, or when aroused by "unbidden" images […], in nightmares or even while awake. The sudden fear–shivering, palpitations, a sense of panic–on the smell of burning in former fire victims is a more common occurrence. Equally common is the unexpected outburst of tears when, many years later, a letter, a toy, a piece of clothing of a child who died is stumbled on; or the blood that rushes to one's face when recalling an embarrassing act committed years ago. (op. cit., 15).

One may suppose that mental reappearances of the eliciting stimuli during emotion episodes are underpinned by the same factors that would underpin the elicitation of new emotion episodes referring to old events, so that unless their causal power is neutralised by habituation, distraction, emotional fatigue and the like, the emoting subjects are susceptible to experience reappearances of the eliciting stimulus which elicit anew emotional experience. In this sense, the hypothesis is compatible with PRO: emotional persistence is explained by continual elicitation occurring because of factors lying beyond agential control. In order to reject PRO, we need to show that the cognitive processes which occur during emotional reactions are not passive cognition of this sort. As we will see in due course, we may do this by showing that emotional cognition is inherently purposive, or more precisely *interest* driven. At first approximation, the aim of emotional cognition is to cope or deal with the demands of relevant emotional situations; as such, is not dependent on mere ideational association. Rather, it depends on an ongoing, *interest*-driven exercise of agential control on the part of

the emoting subjects over their attentional selectivity. What I wish to claim at this stage, though, is simply that the concept of 'reaction' is in principle *compatible* with this sort of view. Undermining the connection between features of the reaction at onset and duration, showing that intensity at onset influences duration by determining the frequency of reappearances of the eliciting stimulus does seem to accomplish that.

The fact that the duration and time-course of emotional reactions may enjoy some independence from their initiatory elicitation, I suggest, is closely linked to their being reactions. Indeed, arguably, the essential property of reactions–what distinguishes them from other kinds of effect–is precisely their distinctive *temporality*: reactions, having been elicited–that is, initiated or kicked-off–take time to run their course. By this I mean that they may outlast their distal elicitors, for the dispositions or sensitivities that underlie their occurrence take time to fully manifest themselves. It is precisely in this respect that reactive aspects of the mind differ from other instances of passivity in the mental domain, namely *receptive* aspects of the mind: their persistence is independent from their initiatory causation. To wit, the persistence of perception necessitates continuous perceptual stimulation, whereas emotional reactions and other kinds of affective reaction only require stimulation to be triggered or kicked off. For example, a pain episode distally caused by my touching a hot stove may and *typically* does outlast my contact with the stove, whereas perceptual experience distally caused by the presence to my visual system of the same stove stops as soon as I turn my visual system or my attention elsewhere.[95] I will call this the 'proper course' of a reaction, since in case a reaction persists in virtue of continued elicitation as opposed to persisting because it is running its course, it may be analysed into a sequence of temporally contiguous reactions which share the same elicitation conditions. In such cases, the reaction is in a sense

[95] So-called 'after-images' may at first be thought to provide a straightforward counterexample, but they cannot be plausibly considered continuations of previous perceptions, i.e., parts of the perceptions' *course* over time, for subject is no longer *perceiving* anything – they are classified as illusions. Rather, they are *themselves* reactions, namely, in the case of visual after-images, retinal reactions to overstimulation of our photoceptors, and hence, to previous perceptions.

renewed at each instant–i.e., at each instant a new punctual reaction takes place. As I have mentioned, the proper course of emotional reactions is closely related to what I have called their 'lifecycle' or natural course of development. It is because they are reactions that emotion episodes have this sort of dynamic character: it belongs to the way in which emotion episodes run their course to undergo changes in intensity outlining rise, peak and subsidence.

Recognising that reactions have proper course is crucial to see in what sense they are passive. Since the concept of 'reaction' implies passivity as regards the *inception* of the phenomenon it applies to, it leaves room for agential factors to enter the scene and determine how the phenomenon will unfold over time as it runs its course. It may belong to its proper course that agential factors must come into play. This nicely explains why one may qualify certain behaviours, from scratching an itch while awake, to virtually any technical movement performed in a boxing match, as reactions; it is not that these are not actions of ours, but that they are initiated by an external stimulus. Analogously, the concept of 'proper course' makes room for agential factors to cause the persistence of reactions in the mental domain. Whether agential factors *do* come into play crucially depends on the nature of the reaction. As I will argue, all the required factors are in place in the case of emotional reactions: our emotional sensitivity–viz., our affective interests–are in fact such that their manifestation involves agential cognitive processes, which essentially contribute to emotional persistence. Failure to see that the course of emotional reactions is a 'point of entry' for mental agency plausibly explains why theories of the causation of emotional particulars equate 'causation' and 'elicitation', and, relatedly, ignore issues of emotional persistence. The persistence of emotional reactions is taken to be dependent on features of the emotional reactions at onset, and, therefore, it is thought of as entirely passive: it is a matter of initiatory elicitation and return to baseline of the elicited changes. This interpretation of the concept of 'reaction' as it applies to emotional reactions is what I have called the Passivist Reaction Ontology (PRO).

In the next section, I will argue that PRO lacks the resources necessary to account for emotional persistence as distinct from other kinds of persistence of affect.

6 Emotional persistence and cognitive mediation

There are just two resources that PRO has at its disposal to account for the duration of affective reactions: (*i*) elicitation and (*ii*) proper course–the latter understood in terms of passive or inertial progression towards subsidence. Elicitation might occur, in principle, (*a*) at inception, (*b*) at inception and then continually, or (*c*) intermittently. It should be noted that elicitation and proper course may combine and mutually interact: the duration of a reaction may be due to its continual elicitation–e.g., tickling-induced laughter–to its punctual initiatory elicitation plus proper course–e.g., pain we feel when someone steps on our toes–or its *intermittent* elicitation in an interplay with proper course–consider, for instance, anger reignited many times during a heated discussion. In all such cases, the continuous occurrence of the reaction is due to the continual obtaining of its elicitation conditions: the (ongoing) impact of specific elicitors and the presence of dispositions or sensitivities, that may take time to fully manifest themselves. Consequently, in specifying the conditions of elicitation of the reactions, one may take oneself to have thereby specified their conditions of *persistence* over time. Such causal structure is quite clearly exemplified by the simpler affective reactions such as pain, hunger or sensory (dis)pleasures. Consider the pleasure one may feel as one enters a warm bath after a long exposure to a much colder temperature. In this sort of example, the distal elicitor of the affective reaction is an increase in ambient temperature. We know that an increase in ambient temperature causally impacts the subject by causing a decrease in the discrepancy between her skin temperature and her core temperature, and that this decrease in discrepancy causes various physiological changes, namely the progressive subsidence of previously ongoing processes of thermoregulation, such as, for example,

shivering and vasoconstriction. To the extent that this sort of changes are felt changes, various bodily feelings or sensations will occur. Now, such feelings are the bases of affective reaction of pleasure: they in turn cause affective changes, namely, positive hedonic feelings. Once we have specified the elicitation conditions of this particular instance of pleasure–namely, the subject's complex disposition to react with pleasure to an increase in the ambient temperature and the change in ambient temperature–in order to account for the way in which such affect unfolds, from initiatory elicitation to the moment in which the subject, having habituated to the new ambient temperature, returns to a neutral affective condition, we can simply invoke the passive course of the affective reaction as ensured by the workings of the subject's thermoregulatory system.

The idea behind PRO is that the same sort of approach is applicable to emotional reactions, for their duration, just like the duration of sense-based and drive-based affective reactions, does not pose any *special* problem, that is, it does not pose problems additional to those posed by the reactions' initiatory elicitation: once the initiatory occurrence of emotional reactions is explained by appealing to certain elicitation conditions, we may also explain their duration by appealing to the continuous obtaining of the specified elicitation conditions. So, can 'elicitation' and a passive, inertial 'proper course' account for emotional persistence in the way in which they account for the duration of other kinds of affective reaction?

Although I maintain that emotion episodes are species of affective reaction, I wish to deny that this classification entails an account of emotional persistence on the model of the persistence of other kinds of affective reaction. The reason is that emotional reactions instantiate a more complex causal structure. This further stratum of complexity is introduced by the fact that emotional reactions are reactions to cognised situations, and their cognitive mediation produces to two structural disanalogies with other kinds of affective reaction. Firstly, the causal structure of emotional reactions, though similar to that of other, *simpler*

affective reactions in that it involves elicitation of affective changes and proper course, differs from it in that the proper course of emotional reactions inherits structural characteristics from the way in which the various agential cognitive processes they involve unfold over time, and, secondly, the relation between affect and cognition is–unlike the one between affect and sensation–of reciprocal causation.

The idea that passive experiential phenomena may *inherit* structural characteristics in this way from mental actions has been introduced in discussions of the nature and structure of ongoing perceptual activities: perceptual activities such as looking, watching and listening share in fact certain structural characteristics with ongoing intellectual activities such as imagining, recollecting and directed exploratory thinking, and, more generally, with instances of sustained attentional engagement with an item or topic. Subjects of perceptual experience associated to ongoing perceptual activities, though *passive* in some essential respects–namely, they *undergo* perceptual experience, and the experience is caused by their being acted on by stimulation of their sensory receptors–are active in other respects: the continuous production of perceptual experience is partially agentially maintained by means of, e.g., looking at, or listening to the sources of stimulation, that is, by an exercise of agential control over the way attention is deployed.[96] Whether an account of emotional persistence in terms of elicitation and passive proper course is right, crucially depends on whether the causal structure of emotion episodes is wholly captured by PRO or shares instead certain characteristics with experiential continuities whose persistence is in part determined by mental agency. I will argue that emotion episodes involve elements of *both* kinds of causal structure. I will then analyse the interplay between passive and agential elements by developing the idea that during emotion affective changes and cognition are *mutually* causally dependent. All of this, as we will see, follows from correct appreciation of the claim that emotion episodes are

[96] See, e.g., O'Shaughnessy (2000); Crowther (2009; 2010).

cognitively mediated. Although this is the main presupposition of broadly cognitive theories of emotion, not much theoretical effort has been devoted to exploring how cognitive mediation may complicate the causal structure of emotion episodes. My main point of departure will be that in so far as the affective changes involved in emotion episodes are dependent on occurrent cognitions as their *bases*, and thus cannot outlast them in the way in which they can outlast their distal elicitors, it is natural to assume that emotion episodes depend in part for their persistence on the continuous occurrence of cognitive processes. Consequently, it is natural to assume that in order to account for emotional persistence, we need to provide an analysis of the causal structure ensuring the continuous occurrence of the bases of emotion episodes.

7 Multi-attitude emotion episodes over time

I am assuming that emotion episodes are reactions to *cognised* concern-relevant situations: their occurrence requires the subject's attention to be engaged or occupied by cognition directed towards concern-relevant situations, i.e., situations that are congruent/incongruent with our concerns. Our standing concerns are what makes us sensitive or responsive to emotional situations, and our taking cognisance of emotional situations is the *proximal* elicitor of the manifestation of such sensitivities, namely the occurrence of emotion episodes. Now, one crucial reason to deny that PRO explains emotional persistence is that it *fails* to do justice to the way in which multi-attitude emotion episodes unfold over time–i.e., by undergoing attitudinal transitions. The reason why it fails is the following: to specify elicitation conditions for emotion episodes requires specifying sensitivities to a given class of elicitors, for instance 'dangerous situations', and elicitors that are members of such class, i.e., whatever it is that the subject takes to be fearsome or dangerous. So, for the elicitation conditions of emotion episodes to remain unchanged relative to an interval of time, as per PRO, both sensitivity

and elicitor must remain constant for that interval. Constancy of both sensitivity and the elicitor would predict that emotion episodes may only persist without any other kind of variation besides variation in intensity: emotion episodes may run their course by going through their rise, peaks, plateaus, and, finally subsidence, but the subject's attitude at onset, e.g., the fear, would steadily endure. If we consider, however, the way in which multi-attitude emotion episodes unfold over time, that is, undergoing attitudinal transitions, it becomes clear that both sensitivities and *proximal* elicitors must vary over time. Imagine an episode of emotion involving recurrent attitudinal transitions amongst 'anger', 'shame' and 'sadness'. Such an episode, if elicited by an interaction with some single and stable situation, for example, a clumsy fall from your bicycle into a big puddle of mud, must clearly involve reappraising the situation under *distinct* evaluative aspects. Let's suppose this situation is relevant to disparate concerns. For instance, we can plausibly attribute to the subject the universal concern for good or smooth functioning (Frijda 2007, 138-141)–a concern posited to explain the anger we feel if ongoing activity is interrupted by some kind of failure–and concerns for strength or elegance or athleticism. Now, for the episode to undergo its changes in attitude it seems that we must grant something along these lines: after an initial burst of anger due to the fall, the subject's attention is sometimes occupied by thoughts that are relevant to the shameful character of the situation–for example, "how frail and clumsy I must have looked falling like that"–and sometimes it is occupied by thoughts that are relevant to its offensive character–e.g., "I *shouldn't* have fallen like that; had I been more athletic I wouldn't have fallen like that." For these different thoughts to occupy the subject's attention and elicit affect different evaluative sensitivities need to be engaged in sequence.

I have chosen an emotion-eliciting situation of little significance to highlight how common multi-attitude emotion episodes are in our occurrent emotional life, and I have considered an episode of emotion starting with a 'burst of anger' triggered by the frustration of our very *primitive* concern for good functioning to introduce another element of the view that I am

articulating, to which I will now turn. Several theorists in psychology have emphasised that what elicits the initiatory occurrence of emotion episodes, is often, if not typically, a rather simple appraisal. Episodes of the most conceptually *sophisticated* kinds of emotion, e.g., 'anger', 'guilt', and 'shame', are often elicited, respectively, by frustration, mere involvement in some harm being inflicted and actual or incipient social rejection (Frijda 1993; Lewis 1995; Robinson 2005). In other words, 'offence', 'culpability' and 'shamefulness' do not always feature in the content of the initiatory appraisal of episodes of 'anger', 'guilt' and 'shame'; rather, they often result from what Frijda (ibid.) calls cognitive *elaborations*. The reason why conceptually sophisticated construals of emotional situations are often taken to be necessary for emotion elicitation, Frijda suggests, is that they feature prominently in the content of our emotional experience; but this does not prove that they must also feature in the content of the initiatory appraisals.[97] If we accept that emotion episodes may have rather elementary proximal elicitors, we may not have to brush under the carpet, as it's often done, some quick counterexamples to the claim that emotions involve *from the start* appraisals of thick values, such as, for example, episodes of anger directed towards inanimate objects or episodes of guilt arising from fortuitous fulfilment of one's wish for someone else's suffering and so on. We may then account for the very high level of complexity that our emotional experience may reach while starting with such elementary initiatory appraisals if we accept that such appraisals arouse not only affect, but also affective interests, which then drive and govern our explorations of the value of the emotional situations–explorations which may over time crystallise in disparate 'assessments' in terms of thick evaluations.

The ongoing process of exploration of value that may underlie the successive engagement of disparate concerns may be viewed as part of the way that emotion episodes *run their course*. As we have seen, one sense in which emotion episodes, like all (affective) reactions, have

[97] Cf. Lazarus' (1991) distinction between primary and secondary appraisal.

proper course has to do with the fact that once they are elicited, they *inertially* proceed towards subsidence: if not causally sustained by continual elicitation processes, moment by moment, their natural course of development is to reach plateau and then decrease in intensity until the subject no longer reacts with affect and starts to return to her neutral affective condition. This way in which emotion episodes may run their course is the way PRO is equipped to accommodate. But the way in which multi-attitude emotion episodes run their course also includes changes in attitude, and this aspect of the course of emotional reactions is one that PRO cannot plausibly accommodate. A dual view of emotion episodes as episodes of affective reaction and episodes of engagement is meant precisely to capture this complexity. The ongoing interplay of affect, interests–viz., emotion attitudes–and sustained attention may plausibly account for the shifts in the aspect under which the situations are appraised as multi-attitude emotion episodes run their course. Over the course of Part II, we will see that such ongoing interplay is a bit more complicated, but no more mysterious than the interplay between interests and attention which takes place in many instances of sustained attention, such as, e.g., instances of problem-solving or directed exploratory thought: in such cases too the interest that drives and governs attention typically evolves, for example, by becoming more and more specific as the search proceeds and begins to deliver its results (cf. e.g., Koralus 2014). In the next two sections, I will prepare the ground for this account by offering an analysis of the causal structure of emotion episodes as distinct from other kinds of affective reaction.

8 Affect, attention, motivation and reciprocal causation

All affective reactions involve motivation: they impel or even compel purposive cognition and behaviour. The motivating dimension of affective reactions is normally regarded as one of their main defining features, a view which goes hand in hand with the very plausible

hypothesis per which the primary biological function of affective reactions is precisely to motivate cognition and behaviour aimed at assessing the personal or biological significance of the affect-eliciting situations and dealing with them, or coping with their 'demands'–e.g., (potential) damages to some body parts in the case of reflexive pain reactions, deficiencies or homeostatic disturbances in the case of thirst and hunger, and concern-relevant situations in the case of emotional reactions.[98] As we will now see, affective reactions arguably owe their motivating *power* to their affective component, i.e., the *inherently* rewarding or punishing hedonic feelings, and not to their *bases*, be them sensations or cognitions–cf. Tomkins (1995); Grahek (2001); Melzack & Casey (1968); Soteriou (2013; 2018).[99] On the other hand, different affective reactions motivate distinct species of purposive behaviour, which raises the issue of how the motivating power of affect obtains this goal-directedness: how does affect motivates us to execute specific actions, varying depending on the affective reaction of which it is part? To capture the specifics of the way in which different affective reactions motivates us to act, I will suggest, we must focus on the relation between affective changes and the bases from which affective reactions derive or *inherit* intentionality. More precisely, I will suggest the way in which affective reactions motivate us has to do with the interplay over time between affective changes, the basis of the affective reaction and the subject's attention, which crucially mediates between the two constituents. To do this, I will first need to argue that some form of ongoing attentional engagement is *constitutively* involved in *all* kinds of affective reactions and that it plays a role in relation to their motivating dimensions.

It has often been noticed that affective reactions consume attentional resources by drawing and fixating our attention onto given affect-eliciting situations, and that this feature of affective reaction is, somehow intuitively, closely related to their power to motivate.[100] Now,

[98] Cf. e.g., Scarantino (2005); Tomkins (1995); Frijda (1986, 2007).
[99] I borrow my characterisation of affective motivation from Tomkins (1995).
[100] De Sousa (1987, 28-30) traces this idea back to Descartes (1649). See Brady (2013, ch.5), Ledoux (1996), Vuilleumier *et al.* (2003). Faucher & Tappolet (2002) and Tappolet (2016, 31-38) provide surveys of the relevant empirical literature and further references.

the relevant intuition, I take it, is that this influence on the way in which our attention is deployed is what imparts a felt *urgency* to the affectively motivated cognitions and behaviour. But we might perhaps articulate this intuition a bit further. To do so, and try to capture the link between attention and felt urgency from the point of view of the subject, I suggest that, while undergoing affective reactions, it seems to the subjects *as if* their capacity to engage in any activity requiring their attention is and will continue to be impaired until they have dealt or coped with the affect-eliciting situations–e.g., until they have protected the potentially damaged body part in the case of reflexive pain reactions, drunk water or eaten some food in the cases of thirst and hunger, delivered a given substance, e.g., nicotine, to their organism in the case of cravings of addiction.[101] In other words, the occurrence of affective reactions forcefully draws attention to the affect-eliciting situations, interrupting or preventing any other activity requiring attentional resource and this crucially contributes to the felt urgency distinctive of such phenomena.[102] In our sense-based and drive-based affective reactions, attention is directed to the somatic feelings associated to the relevant physical condition–e.g., some bodily damage or some organismic deficiency–if not continuously, for our capacity to effortfully redirect attention is, to an extent, preserved, at least continually or intermittently. The *intensity* or persistence of our ongoing attentional engagement with given affect-eliciting situations, and, conversely, the difficulty we may encounter in attempting to disengage our attention and devote it to other activities are arguably positively correlated to the intensity of the affective component: the more intensely we are affected, the more our engagement with the affect-eliciting situations becomes non-optional or compelled. This plausibly explains

[101] Nash (1989) develops a different proposal about the nature of the relation between attention in emotion and urgency: he elaborates on the idea is that having our attention passively focussed onto emotional situations generates *overevaluations* of their significance, hence, the felt urgency of the emotionally motivated behaviours. The idea that affect involves systematic overevaluation does not have much initial plausibility in connection to sense-based and drive-based affective reactions and may only convince in connection to a small subset of emotional reactions, as Nash himself concedes.

[102] Focus on this potentially disruptive feature of affective reactions and neglect of its productive counterpart, i.e., facilitation of affectively motivated behaviour, is what plausibly explains the traditional treatment of emotions as disorganised and disorganising perturbations: Cf. Leeper (1949) for survey and early criticism of this once standard treatment of emotion in empirical cognitive psychology.

why behaviour motivated by intense pain or hunger often takes precedence over other courses of action, even if we have good reasons to pursue other goals, such as for example, winning the match despite the pain, or continuing to work despite the hunger: if the affective reaction is strong enough, we cannot devote attention to the cognitions and actions required to pursue those goals until the demands of the affect-eliciting situation have been answered.

Now, if all affective reactions as inherently motivating phenomena must involve engagement with affect-eliciting situations, as I have suggested, in what does engagement characterising emotion episodes differ from engagement characterising the sense-based and drive-based reactions? The crucial difference, I will suggest, regards whether our engagement contributes to determining the persistence of the affective reaction, that is, whether it is an element of their causal structure. As we will now see, that sustained attentional engagement is an element of the causal structure of emotions episodes is actually implied by their cognitive mediation. Recall that the claim that emotion episodes are caused or causally mediated by cognition is the claim that if an event, object or state of affairs is to cause an emotion episode, then, such an event, object or state of affairs must somehow be cognised by the subject, that is, it must be noticed, imagined, remembered, and more generally, it must occupy or engage attention. It may seem to follow from such characterisation of the relation between emotion episodes and cognition that the first depend unilaterally on the second. Mulligan (1998) problematises such a view of the basing relation obtaining in emotion in the passage I quote below.

> There is a popular view according to which cognitive and perceptual bases cause emotional episodes. The sight of the dog causes fear. This view is certainly true of some situations. It is compatible with the claim that there is an internal relation between the very same terms of the causal relation. But it is a great simplification. First of all, because many perceptions and thoughts which come to function as bases for emotions are themselves determined by other emotions and emotional sensations. Secondly, and more generally, because bases and the emotions based on them often have common causes. And,

> finally, because there are relations of *interdependence*, causal and non-causal, between a person's affective and cognitive/perceptual traits. Ulric's feelings, at a time, for and about his sister are based on thoughts that themselves owe their existence to earlier emotional reactions. (op. cit., 93 italics added)

In what follows, I will push some of Mulligan's remarks one step further, towards a *nonlinear* model of the causal and the internal relations between affective changes and cognition. Mulligan suggests that cognitive presuppositions of our emotional reactions may themselves be *caused* by other (previous) emotion reactions.[103] My suggestion will be that affective changes and cognitions as parts of emotion episodes are *always* related by mutual causation.[104]

Let me begin by noting that the *simplistic* account of the basing relation, does not sit well with striking phenomenal characteristics of emotional experience. During emotional experience, at least in those cases in which we have awareness of our thinking processes, it often seems to us that our emotional reactions persist or continue precisely because we are engaged in certain thinking processes, for example, it may seem to us that our fear or our anger will not subside because we do not stop thinking of fear- or anger- eliciting situations. Nonetheless, it may also at times seem to us that we persist in certain thinking processes precisely because we are currently in the grip of emotion. Both directions of causation are reflected in our folk-psychological explanatory frameworks: emotions are often invoked to explain certain thoughts and thoughts are invoked to explain emotions. For example, we may say that we cannot stop thinking about what our partner might be doing because we are in the grip of jealousy, but also that we keep suffering pangs of jealousy because we just keep thinking about what our partner might be doing. This shows that the view of the relation between emotion episodes and their cognitive bases on which the first are unilaterally dependent on

[103] See also Mulligan (2007). Goldie's (2012) account of the dynamic relation between emotions and moods he develops drawing on Robert Musil's novel *The man without qualities / From the posthumous papers* (1995, 1239-1311) is also highly pertinent.
[104] Versions of this view are explicitly defended in psychology by Frijda (1993) and Lewis (1996) among others.

the second simplifies the phenomenon as Mulligan suggests. Although in some cases we do seem to take the emotion episodes we undergo to be caused by our thinking processes, in other cases, we seem to take the emotion episodes to be the causes of our thinking processes. Relatedly, there seems to be in emotional experience a mixture of passivity and activity (Goldie 2000): emoting subjects feel acted upon in their being affected by given thoughts which just keep occurring to them, but they also feel active in their (impelled) engagement in affect-eliciting cognitive processes such as ongoing imagining, elaborating, ruminating. How can we reconcile these two directions of causation? This is the question I will address. My main proposal will be the relation holding between affective changes and cognitions, unlike the one between affective changes and sensations, is one of reciprocal causation: emotion episodes exemplify a distinctive causal structure crucially involving two distinct but closely related agencies: elicitation of affective changes and sustained emotional engagement. Recognising the involvement of both causal agencies as determinants of the occurrence and persistence of emotion episodes is crucial if we are to correctly theorise the relation between affective and cognitive elements in emotion episodes and relatedly the mixture of passivity and agentiality seemingly characteristic of emotional experience as distinct from other kinds of affective experience. In what remains of this chapter, I will argue for this proposal by highlighting crucial *structural* contrasts between emotion episodes and sense- or drive-based affective reactions. The main upshot of the analysis will be that in order to account for the reciprocal causation at play in emotion episodes we need to appeal to the notion of *interest*.

9 Attention in sense-based and drive-based affection *vs* emotional reactions

I will start with an analysis of the structure of pain, or more precisely reflexive pain reactions. I will thus exclude, for example, chronic pains. The analysis draws heavily on Grahek's (2001) and Soteriou's (2013) discussions of such phenomena.

We can start with the very plausible observation that subjects of located painful sensations are–and *feel* that they are–impelled or *compelled* to attend to such sensations in a way in which they would not if the latter weren't *painful*: located painful sensations are, typically, difficult to ignore; they are more or less obtrusive, depending on their extent, quality and intensity, how they compare in these respects with other physical sensations with which they compete for our attentional resources, and how absorbed we are in ongoing (mental) activity. Now, one might suppose that what compels our attention is the inherent salience of the located sensations. However, there's strong empirical evidence that what draws attention to such sensations is not their extent, quality, or intensity *qua* located sensations–i.e., their inherent sensory salience–rather, it is their 'painfulness' understood as an affective element that may be distinguished from sensory phenomenology. Consideration of 'dissociative syndromes' in pain experience, and, especially, pain asymbolia[105] seems to show that conscious awareness of given bodily located sensations–what Grahek (2001) aptly calls 'sensory-discriminative' components of the pain experience–in the absence of an aversive *affective* reaction to their occurrence "doesn't move the […] mind in any way and is all too easily ignored" (72-73). On the other hand, when one is in a state of pain *because* of the occurrence of located painful sensations, one's attention is not occupied by an unlocated suffering–as it may be the case, for example, in exhaustion, emotional suffering, or unlocated unpleasant sensations such as the ones that are sometimes generated by given tastes, sounds, or colours.[106] For the suffering involved in reflexive pain reactions to occupy one's attention is for one to have one's attention occupied by particular sensations located in given parts of one's body, sensations that have spatial and temporal features, and that typically instantiate distinctively 'painful' qualities, such as, for example, being 'cutting,' 'drilling,' 'burning,' 'pulsating,' 'throbbing' etc. (cf. e.g., the *Mc Gill pain index*).[107] This point is clearly highlighted by another dissociative

[105] Grahek (2001). Massin (2020) and Soteriou (2013, §3.5) for diverging interpretations of Grahek's discussion.
[106] Cf. Sherrington (1900, 968-967).
[107] See Melzack (1975).

syndrome in the human experience of pain that is also discussed by Grahek. This syndrome is characterised by "loss of pain sensation with preserved pain *affect*" (Ploner *et al.* 1999, 211), and may thus be viewed as the converse of pain asymbolia: although the patients are *in pain*, they are unable to (exactly) spatially localise the painful sensations, to (exactly) register their temporal properties–e.g., their onset and termination–to evaluate their intensity, as well as to qualify the sensation more precisely than as generally unpleasant–i.e., they cannot say if the sensation is, 'cutting', 'burning', 'throbbing' etc. (Grahek 2001, 100-104).

Consideration of the two dissociative syndromes suggests a way of theorising the respective and distinctive contributions of the affective and the somato-sensory components of pain: the former seems to contribute the negative valence, or aversive character of the reaction, and the latter seems to contribute information about its spatial and temporal features, sensorial intensity, and 'painful' qualities.[108] In light of the analysis of the relation between affective motivation and attention that I have put forward in the previous section, I suggest that while affect determines the persistence of attentional engagement with affect-eliciting situations, the somato-sensory component determines the orientation of attention, that is, its target or focus. This double modulation of the way in which attention is deployed in affective reactions confers urgency to the affectively motivated behaviours, but also, crucially, *specificity*: during affective reactions it feels to the subjects as if their capacity to engage in any other activity requiring attentional resources will continue to be impaired until they have dealt with the *specific* demands of the affect-eliciting situations in *specific* manners. To illustrate, both reflexive pain reactions and drive-based affective reactions, such as, e.g., hunger, involve felt aversion directed towards somatic conditions, but, in virtue of the internal relation between such somatic conditions and our suffering, or discomfort, as

[108] Cf. Massin (2020) for defence of the view that pain and suffering are distinct phenomena, and more precisely, that *suffering* is an *attitude* and the correct reaction to pain. This proposal is analogous to Soteriou's in that it recognises the distinction between the somatic aspects and the affective aspects of the phenomenon. However, it crucially differs from it in that it doesn't see these aspects as components of some single complex reaction.

mediated by attention, we do not feel motivated to make our suffering or discomfort cease by any means available. Rather, we feel motivated to engage in specific actions that feel to us to alleviate the situation. For example, reflexive pain reactions motivate us to protect the body part where the painful sensation is located–and not to take pain killers–while hunger motivates us to seek food and eat–and not to take appetite suppressants.[109] The point is nicely elaborated by Soteriou (2013), who provides a detailed analysis of pain with special emphasis on the nature of the internal relation holding between the affective and the somato-sensory component.

> From the subject's point of view, it is not simply as though some event occurring in some part of her body is causing her to be in pain, where an event of the same kind could in principle occur in that body part without its having that effect. The kind of located sensation-event that is occurring feels to the subject to be one that couldn't occur without her being aware of it and without her being in pain–without that part of her body hurting. So, it feels to the subject as though it is not possible to stop the unpleasant effect of the occurrence of the located sensation-event without stopping the occurrence of that located sensation-event. *Therefore, the subject won't simply be inclined to avoid the unpleasant effects of the occurrence of the sensation-event by protecting herself from the located sensation-event.* The subject's natural inclination will be to protect herself from whatever it is that is causing the occurrence of the located sensation-event, and this will often require protecting the body part where she feels the sensation-event to occur. (op. cit. 75-76, italics added).

In this passage, Soteriou stresses that when we undergo pain reactions, the painful sensations, and our state of 'being in pain' are not felt by us to be related merely as cause and effect. Rather, the relation between the subject's being in pain and the located sensations is one of 'constitutive dependence', a dependence which is reflected, as he suggests, in the explanatory

[109] Of course, when in pain or hungry we may acquire motivations to take pain killers or appetite suppressants. But in such cases the pain or hunger motivate by giving us practical reasons (see Ombrato & Phillips 2020), and this sort of motivation, as I will argue later on, is crucially different from affective motivation (ch.6).

circularity of any complete specification of the nature of the painful sensation and subject's suffering or state of *being* in pain: this particular state of being in pain is in part to be specified in terms of the occurrence of the particular localised sensation in virtue of which it obtains, and the localised sensation is in part to be specified in terms of the obtaining of the state of *being* in pain (ibid.). The internal relation between somato-sensory and affective components explains why during pain we feel an urge to protect the body part where the painful sensations are localised, whereas, in hunger, for example, we feel an urge to eat, that is, to remedy the situation that is causing the occurrence of the unpleasant sensations of hunger, and not, crucially, an urge to seek to rid ourselves of such unpleasant sensations by any means available.

Let me now turn to emotion episodes. The structure of emotion episodes is analogous in all relevant respects, and, crucially, in respect to the mediating function that our attention plays. In both kinds of affective reaction, the subjects' attention is engaged by the affect-eliciting situations and the engagement is felt by the subject to be impelled or coerced in the sense clarified. Moreover, the emoting subjects do not seem to take the various cognitive processes that provide the episodes with their bases *merely* as causes of episodes.[110] Rather, they take such mental activity to be such that they couldn't be engaged in it–in that particular moment–without being 'angry', 'joyful', 'ashamed', 'relieved', 'resentful' etc. In other words, during emotion we tend to take our thoughts to be inherently 'pleasurable' or 'unpleasurable' in such determinate ways, and, relatedly, it feels to us as though the negative or positive hedonic feelings we enjoy cannot be stopped as long as such thoughts continue to occur to us. Accordingly, we are not typically motivated to make our emotional sufferings stop by any

[110] This intuition is less compelling in the case of emotion episodes as compared to the pain case: we do sometimes take out thoughts to be the causes of our emotions. However, this seems in part as product of our relatively sophisticated understanding of these phenomena. From a naïve subjective point of view, arguably, our emotional reactions are felt to be caused by emotional situations.

means available; rather, we are motivated to make it stop by trying to cope or deal with the demands of the emotional situations that we take to cause our mental activity.

The fact that emotion episodes have their bases in cognitions, however, imports a number of differences, stemming, ultimately, from disanalogies between perceptual and intellectual (sustained) attending. Firstly, our sustained attentional engagement with emotional situations and the continuous occurrence of emotion-eliciting cognitions are one and the same thing: our being engaged with emotional situations is not separable from attention being occupied by emotion-eliciting thinking processes. This is due to the fact, remarked by Peacocke (1998), among others, that our thinking processes are not, properly speaking, 'objects' of experience; we do not *orient* our attention *towards* thinking processes, as if they were there independently, or prior to, their occupying our attentional resources–i.e., as parts of some cognitive analogue of our field of perception, i.e., *objects* of which we may have some pre-attentive awareness.[111] The engagement with located painful sensations and the sensations themselves, by contrast, are separable occurrences.[112] Of course, the way our attention is deployed during episodes of pain might affect–for instance, intensify–our awareness of the painful sensations, and, hence, our being in pain, but this familiar psychological fact, if it is correctly theorised, confirms that our enjoying bodily located sensations and our being attentively aware of them are not

[111] This important disanalogy between the way that attention relates to cognition and sensation or perception is nicely illustrated by Martin (1997) as follows. "When I notice at the periphery a slight movement, and in scanning the foliage fix briefly on a flash of scarlet, my attention is drawn to a new area of the visual scene and I am impelled to find out more about what has drawn my attention in this way. In this kind of case, it is tempting to think of experience in terms of a whole array of items stretching beyond what I have focused my attention on at a time […]. It is as if I am aware of the whole array at a time, […] and my awareness of some element of it can explain why I shift my attention from one part of the scene to another. *There seems to be no corresponding array of items to shift one's attention over in thought*: if we think of thoughts as determinations of attention, then there can be no way of thinking of something without thereby to some extent to be attending to it. *If we are not currently sensorily aware of something, then if we are not thinking of it, we do not take ourselves to be aware of it*." (78, italics added). Cf. McGynn's (2004, ch.1) discussion of the 'attention-dependence' of thought and mental images.

[112] I am here assuming that attention is *not* necessary for consciousness, and, hence, the correctness of '2-level' theories of consciousness (Lambie & Marcel 2002), namely, theories which identify two forms or modes of consciousness: (*i*) a *pre*-attentive, phenomenal consciousness and (*ii*) an attentive, recognitional consciousness. To hold firm this distinction is to hold that, at any given time, one is conscious of more than that which currently occupies one's attention. See, e.g., Dainton (2000, ch.2), O'Shaughnessy (2000, chs.17-18), Searle (1992, ch.6). Intriligator & Cavenagh (2001) and Marcel (1993) provide empirical evidence. Wu 2014 (chs.5-6) offers a useful framing of the debate, evaluation of some of the empirical evidence and further references.

one and the same thing. As O'Shaugnessy (2000, 489) observes, our attention may intensify our *awareness* of the sensations, not the sensations themselves, and this precisely because the sensations are available to us independently of their occupation of attentional resources.[113] This is why our attentional engagement doesn't contribute to the persistence of the simpler affective reactions: the affect-eliciting sensations, though *mind-dependent* phenomena, are not 'attention-dependent'.[114] And it is why it *does*, crucially, contribute to emotional persistence: in absence of sustained emotional engagement, the cognition which provides the emotion episodes with their bases would cease to occur and the emotion episodes would therefore terminate. In other words, our being *engaged* with emotional situations consists exactly in our having certain images, thoughts, and memories, namely the images, thoughts, and memories by which we happen to be currently affected. This observation leads us to another crucial difference, a difference, which, I will argue, confirms and specifies the intuition that emotion is closely related to 'interest.'

I have suggested that in drive-based and sense-based affective reactions the affective changes determine the *persistence* of our attentional engagement with affect eliciting situations whereas the somato-sensory component determines its orientation. These changes in the way our attention is deployed, underpin, respectively, the impelled or coerced character and the specificity of affectively motivated behaviour. Analogously, occurrent cognitive processes, since they are the bases of emotion episodes, might be thought to determine the selectivity of the occupation of emotional attention–and, hence, the specificity of emotional motivation. The reason why one may call upon the somato-sensory component of drive- and sense-based

[113] It is worth quoting the relevant passage in full. "We distinguish *faint* awareness of *intense* sensation from *intense* awareness of *faint* sensation. The governing principle is, that degree of awareness of a sensation is distinct from the degree of the sensation. Therefore the degree of attention directed onto a sensation need have nothing causally or logically to do with the intensity of the sensation. A faint or marginal awareness of an intense sensation is thus a genuine possibility: a sensation need not diminish in intensity as a result of the poverty of attention coming its way. Then why not the occasional total unawareness of some sensation? Or are we to believe that, as a medium-intensity sensation passes from faint almost subliminal awareness into total unawareness, it automatically topples off the ledge into non-existence? I can see no justification for this view." (O'Shaughnessy 2000, 489).

[114] I borrow this terminology from McGynn (2004, ch.1).

affective reactions to play this role, however, is that sensations are *not* attention-dependent. Our *pre*-attentive awareness of them is what makes room for the explanatory connection; that is, it creates the *gap* between the occurrence of the sensations and attention-orienting: when in pain, the located sensations–with their spatial extent, quality and intensity–may *capture* and retain our attention, and the persistent attentional engagement with them that may follow–e.g., if the pain is too intense for us to re-orient our attention–is not sustained by us: it consists in passively tracking, moment by moment, their evolution over time. Now, in cognitive contexts, as we have seen, there is no gap; the mental images and thoughts that we enjoy during emotion *cannot* be taken to drive the attention-orienting. What might then play this function in emotion episodes? This question concerns what I will call the *locus* of control of attention-orienting (Posner 1980), i.e., the property of the (orienting of) attention of being exogenous–i.e., *stimulus*-driven–*vs* endogenous–i.e., *purpose*-driven (cf. Posner 1980). While in the 'pain' case and in the case of other kinds of sense-based and drive-based reaction the attention-orienting is controlled by the bases of the reactions–i.e., the located sensations–in emotion episodes, the bases of the reaction– the cognitive processes themselves–are not fit to play this role. The fact that during emotion we feel as if mental images and thoughts associated to our current evaluation keep occurring to us does not mean that they *draw* our attention onto them as sensations do: since we have no cognitive analogue of the perceptual field, the selective occupation of emotional attention by emotion-eliciting cognition is *not* plausibly controlled exogenously.[115] Now, to make sense of continuous selective occupation

[115] The same holds for cases in which the emoting subjects are perceiving the emotional situations. I will discuss emotional attention in sensorial contexts in more detail in the next chapter, but the key idea at this stage is that while bodily located painful sensations cause affective changes immediately, perceptions of emotional situations cause emotion episodes only in so far as the subject notices, or 'is struck by', an *aspect* of such situations–e.g., their 'horrific', 'fearsome' or 'disgusting' character–and then remains *cognisant* of them as evaluatively connotated, that which may involve perceptual-evaluative exploration–cf. Neisser's (1976) cycles of perception. On the other hand, we are not tempted to describe a situation in which our attention is continually drawn to some painful sensation as a situation in which our being in pain has aroused an interest in how our body feels. There are some cases in which it might by appropriate to say that pain has aroused an interest, but these are cases in which it is no longer *mere* pain, or thirst, or hunger that we are feeling, but pain, or thirst, or hunger plus some emotional experience–e.g., pain is scary, thirst and hunger are often distressing. In other words, the relevant units of analysis outstrip the synchronic boundary of the simple affective reaction. Cf. Tomkins (1994).

of attention such as that we observe in emotion it is natural to invoke the notion of 'interest': whenever someone's thoughts succeed one another *intelligibly*, i.e., in accordance, apparently, to a principle of selection, it is natural to suppose that that person is interested in something, and that her current interest imposes specifiable constraints on what may and what may not happen to occupy her attention moment by moment. This suggests that we can account for the continuous selective occupation of emotional attention by emotion-eliciting cognition by building active states of interest into the emotion episodes. To clarify, I am not proposing that we introduce some further emotional component; the various specific emotion attitudes understood as an attitudinal-evaluative component of emotion episodes are fit to play this role of active states of interest. I am elaborating Frijda's (2007) suggestion that emotions are 'variants of interest' as the basic form of action readiness: emotion attitudes are ways in which our interests in emotional situations as concern-relevant or somehow significant crystallise into more specific interests. Elaborating this suggestion, we can put forward the following hypothesis: the initiatory elicitation of affective changes arouses in the emoting subjects an interest in the eliciting situation as evaluatively connotated–e.g., an interest in the situation as 'offensive', or 'fearsome', or just *bad* for one. This affect-aroused interest is then kept alive by affect, and, for as long as it is, it constrains the way in which the emoting subjects exercise agential control over their attention. For example, the initiatory elicitation of an episode of fear will arouse in the subject an interest in the eliciting situation as fearsome, that which can explain why fear episodes involve the occurrence of mental images and thought bearing, say, on the gravity of the threat, the extent of one's vulnerability, one's potential to cope with it. The emoting subject's *interest*-driven mental activity, depending on how it evolves over time, may then specify, preserve, or radically change the original affective interest, and, hence, the affective reaction. These disparate scenarios can explain the varying diachronic complexity

See Ombrato & Phillips (2020) for an account of the relation sense- and drive-based affective reaction bear to such emotional reactions in terms of 'recruitment' and for further references.

of emotion episodes, which may range from the *simple* single-attitude emotion episodes to more and more articulate multi-attitude emotion episodes.

Let me summarise the discussion so far and take stock. The central upshot of my analysis was that the causal relation between affective changes and cognition, unlike the causal relation of affect and sensation, is one of reciprocal causation: for the emotion to persist we need both continued elicitation of affective changes and that the affective changes continue to produce the emotion-eliciting cognition. The continued elicitation of affective changes was shown to causally depend on sustained attentional engagement; this is one of the two discernible causal processes that are required for emotion episodes to persist and unfold; causation goes here from occurrent cognition to affect. What then remains to be examined is the causal process which goes in the other direction of causation: from affect to cognition, and more precisely, from affect to the continuous *selective* occupation of the emoting subjects' attention by emotion-eliciting cognition. Now, such selective continuities usually reveal an underlying continuity of interest, that which naturally suggests that we can try to account for their occurrence by building affective interests into emotion episodes as their components. In the rest of this thesis, I will thus focus on the way in which interests, and more generally, 'purposes', may drive and govern our attention and on how affective interests differ from other mental purposes. Before I move onto this task, I will raise the question whether and how the analysis put forward bears on the issue of how we should understand the emotional subjective point of view.

Conclusion

It seems that one could grant the structural differences between sense-based and drive-based affective reaction and the emotional reactions while downplaying their significance. One may insist in fact that highlighting such differences does not take us any further if the aim is to

identify and theorise the *emotional* in emotional experience, for that inheres in what emotion episodes share with other affects: an impelled or coerced character of attention-orienting as an element of passivity. In other words, one may object to the significance of the hypothesis that sustained emotional engagement is *interest*-driven by insisting that in so far as emotion involves a felt loss of control over attention-orienting, our sustained emotional engagement is nonetheless–in an arguably more relevant sense–exogenous. This idea might be elaborated by proposing that if it is true that mental images and thoughts cannot strictly be taken to *capture* and retain our attention as sensations do, we must then introduce a further distinction, that separates our autonomous and our heteronomous mental *activity* so as to restrict the (significantly) endogenous to the former. There is a worry behind this thought that we may express as follows: if we build active states of interest into emotion episodes, we end up construing emotional cognition too much like we construe deliberate cognition. In one sense, this worry misses the point of my analysis, since it does not give due weight to the affective aetiology of the driving interests, and, relatedly, the impelled or coerced character that our engagement will derive from being driven by such interests. Wollheim (1999, 117) discusses a worry very close to the one at issue, namely that if we make room for 'thought' in emotion we unacceptably "intellectualise" the phenomena. He puts it to rest is as follows:

> This argument erroneously assumes that, inside emotion, thought will operate in the same way as it does inside, say, inquiry, and it overlooks the fact […] that thought is a merely instrumental disposition. Thought takes on an end from the outside. So, when thinking is made to serve inquiry, it serves the end that inquiry pursues: it aids in the construction, or purification, of some truth-oriented picture of the world. Equally, *when thinking is recruited into the service of emotion, it helps to strengthen, or elaborate upon, some attitude that we have towards something in, or held to be in, the world.* It follows that, if thinking intellectualizes belief, there is no reason to conclude that it will intellectualize emotion. When Othello entertains the thought that Cassio possesses the handkerchief spotted with strawberries, and, under the tutelage of

> Iago, turns it over and over in his mind, he does not intellectualize his jealousy. He plunges deeper and deeper into the jungle of his imagination, as he is the first to realize. (ibid., italics added).

In pretty much the same spirit, I have suggested that sustained emotional engagement, in so far as it is governed by interests which originate in affect, is a special, *affective* variety of purposive attending–viz., it is purposive attending "recruited into the service of emotion." Building affective interests into emotion episodes allows us to keep in sight the structural difference between emotion and other kinds of affective reaction while doing justice to the fact that emotion often involves a felt loss of control. On the one hand, we are not tempted to describe a situation in which our attention is continually drawn to some painful sensation as a situation in which our being in pain has aroused an interest in the way our body feels. On the other, affective interests, due to the *urgency* they inherit from their affective aetiology, tend to take precedence over other motives we harbour, and may clash with the ones that we aim to prioritise at the moment–consider, for instance, how our anger may interfere with the aim to defend ourselves from offensive accusations clearly and dispassionately. Plausibly, the more we *strive* to prioritise our deliberate purposes over our active states of affective interest, the more we experience the emotional cognition as 'intrusive', the more we feel that we lack control on our attention.

Insisting on exogenous characters of emotional attention, however, is extremely common. Indeed, in affective sciences, emotional attention is theorised as standing in opposition to purposive attention and cognitive control functions, and philosophers working on emotion are often very keen on exploiting this opposition–for example, in discussions of emotional *recalcitrance*.[116] Now, although the analysis I offered has the theoretical resources to do justice to the fact that emotion often involves a felt loss of control on attention-orienting, there's a

[116] Cf. Vuilleumier (2005); Pessoa (2013). On recalcitrance see, e.g., Brady (2013, §5.1), Tappolet (2016 §1.5). cf. Nash (1989). This conception of emotional attention is discussed in more detail in the next chapter.

sense in which to insist on exogenous characters of emotional attention is to press us to spell out the 'experiential' (i.e., phenomenological) significance of an approach which builds states of interest *into* the emotions episodes. The more common alternative is in fact to understand the felt loss of control on attention-orienting in terms of *two* potentially conflicting processes: (*i*) emotional attending conceived of as exogenous; (*i*) an endogenous cognitive engagement with emotional situations occurring alongside emotional reactions–i.e., an ongoing process of conscious (re)appraisal.[117] The pressure, therefore, is to illustrate the significance of the proposed ontological approach in relation to the question of how to best account for the emoting subject's point of view on their own experience. The broad aim of Part II will be precisely to do so. Let me now conclude Part I by quickly illustrating what is at stake here and why the proposed ontological approach might be taken to bear on the question of the nature of the emoting subject's point of view. It is helpful to this end to contrast my account of the relations between emotion and attention with the one put forward by Brady (2013). Brady, who shares the above worry, elucidates it in reference to the idea that, if emotions may serve certain epistemic functions in connection to the evaluative dimension of our life, it is precisely because emotional attention is "passive, reflexive, non-voluntary and effortless" (op. cit., 219). As we will see, Brady's insistence on these features is motivated by the fact that he sees the epistemic functions of emotion as being a matter of 'alerting' or 'signalling': the initial occurrence of emotion *alerts* us to the presence of concern-relevant objects/events, and emotional persistence is what ensures insistent signalling. Brady is keen to emphasise both emotional persistence and attentional persistence, but on his view, crucially, the former is taken to be prior to the latter in the order of causal explanation: it draws and keeps fixed emotional attention onto emotional situations, whereas any agential cognitive engagement with such situation is taken to occur *outside* the emotion–i.e., it is an endogenous investigation

[117] Cf. e.g., Gross, ed., (2007); Brady (2013, ch.5)

aimed to assess if one's initial emotional reactions are in fact justified.[118] The general picture of the emoting subjects' point of view on their own experience that may result from Brady's treatment of the relations between emotions and attention is therefore very different from the one that may result from an analysis of the structure of emotion episodes such as the one I have proposed. On Brady's approach, in so far as the emoting subjects may actively engage with emotional situations as objects of interest, engagement is driven by deliberate purposes. Over Part II of this thesis, I will build up to an account of the way in which mental agency is exercised in emotion episodes aimed to show that this sort of approach seriously misrepresents the emotional subjective point of view.

[118] Cf. Brady (op. cit.): "On this picture, emotional persistence can enable the subject to determine whether what appears to be dangerous or shameful really is dangerous or shameful, through keeping attention fixed on these questions and promoting critical reflection about them". (93).

PART II

Chapter 4
ATTENTION, SELECTION AND TIME

Introduction

Over Part I of the thesis, I defended an approach to emotions aiming to do justice to crucial ontological properties of such phenomena–i.e., their complex, episodic, dynamic character. To do so, I specified a criterion of individuation for (synchronically and diachronically) complex episodes of emotion as opposed to episodes of emotion identified in reference to continual attitude retention. I then asked which single causal process or mode of production may account for the way in which such episodes persist or unfold over time. My answer was that emotion episodes, unlike *simpler*, sense-based and drive-based kinds of affective reaction, are caused to persist as they do by an ongoing interplay between passive and agential factors–i.e., affective changes and our sustained attentional engagement with emotional situations. One of the reasons why this approach is attractive, or so I suggested, is that by letting us keep in sight the passive and the agential aspects of emotion as well as the ongoing interplay between the two, it may allow us to attain a better understanding of the emotional subjective point of view. Over Part II of this thesis, as anticipated, I will substantiate such suggestions. This will involve specifying how mental agency is exercised in emotions and identifying the distinctive *self-awareness* associated to such an exercise of mental agency. My main working hypothesis, as also anticipated, will be that sustained emotional engagement is an instance of

purposive, sustained attending: all forms of sustained attentional engagement with an item or topic, as distinct from what I will call 'idle' ongoing attending–e.g., 'mind-wandering' or 'daydreaming'–are in fact characterised by continuous selective occupation of the subject's attentional resources. My strategy will be to first explore this more general phenomenon and then progressively differentiate emotional engagement and unemotional engagement. Let me outline the broad structure of my argument.

In Chapters 4 and 5, I shall mostly leave emotion aside and focus on (sustained) attention. This will allow me to further clarify the significance of the diachronic point of view in reference to mental agency. Chapter 4 is preliminary and introductory. Its key aim is twofold: to secure the claim that sustained emotional engagement is endogenous by addressing the objection that was raised at the end of the last chapter and prepare the ground for the identification of the respect in which mental agency is exercised in (sustained) attention by clarifying the relations between mental agency and attentional selectivity. The main upshot will be that mental agency in attention is exercises *essentially* over time; it is agential control over what I will call diachronic selectivity as opposed to synchronic selectivity or 'shifts' of attention. The aim of chapter 5 will then be to locate mental agency in (sustained) attention as opposed to idle ongoing attending and characterise the distinctive *self-awareness* associated to this 'activeness.' At the end of the chapter, I'll go back to sustained emotional engagement and specify two sub-classes of attending to which the latter arguably belongs. In other words, I will propose two insufficient, but arguably necessary properties of emotional attending. Anticipating a bit, the members of the two sub-classes are characterised, respectively, by exploratory 'open-endedness'–i.e., lack of completion conditions–and 'non-instrumentality': the drive keeping them on their course by preventing distraction is related to a certain degree of absorption *vs* an awareness of their likely instrumentality to achieving certain purposes. I'll then relate the kind of 'absorption' that we find in emotion by relating it to what some psychologists like to call 'intrinsic' interest; sustained emotional engagement is *interest* driven

in this sense, so I will argue. Differentiating this phenomenon from unemotional absorptions will be the aim of chapter 6. To do so, I'll explore the relations between positive and negative affective interests–understood as kinds of intrinsic interest–and positive and negative value. Drawing on Johnston (2001), I will argue that sustained emotional engagement is motivated by a distinctive, emotional awareness of value, as opposed to an 'intellectual' awareness as philosophers like Brady (2013) seem to hold. Now that I have outlined the broad structure of Part II of this thesis, let me quickly introduce the main argument of the present chapter.

I concluded the last chapter by sketching an objection to the view that emotional attention is endogenous. The objection ran as follows: even accepting that emotions differ from other kinds of affective reaction in structural respects, and that accounting for such differences may lead us to build active states of interest into the causal structure of emotion episodes, emotional attention is still, in the most relevant sense of the term, exogenous. My first task will be to identify the rationale for this objection. My suggestion will be that this rationale is closely related to the supposed centrality, as paradigms of emotions, of emotion episodes that are triggered by *perceptions* of emotional situations. I will try to undermine this rationale by suggesting that all varieties of sustained attention–as opposed to 'shifts' of attention–cognitive and sensorial, are endogenous, and showing that this fact does not eliminate the passivity the objector requires of emotional attending–since such passivity pertains precisely to shifts of attention at an instant. Secondly, I will turn to the link between 'attention' and 'selection', and I will propose that we can distinguish two *ways* in which our minds manifest attentional selectivity: 'synchronic' and 'diachronic.' I will argue that locating mental agency in attention requires articulating the relations between the two. Before I move onto my task, a terminological clarification seem in order. The idea that sustained emotional engagement is purposive was introduced by highlighting the continuity of 'interest' it appears to manifest. However, in discussing the more general phenomenon of sustained attention, I will use the term 'purpose' and not 'interest', so as to save the term 'interest' for the purposiveness of

interest-driven sustained attention. The concept of 'purpose' is in fact neutral as regards the motivational phenomena to which different instances of sustained attentional engagement may owe their purposiveness: it applies to voluntarily prioritised intentions, intrinsic interest, sense-based and drive-based affective reactions, emotional reactions etc.

1 Perceptual attending, purposes, and salience

As I have mentioned, the idea that emotional attention is essentially exogenous is widespread. Indeed, in affective sciences, emotional attention is almost always categorised as 'exogenous', and there is vast empirical (i.e., neuroscientific) evidence which seems to back up such categorisation–i.e., the neural circuit which underpins 'emotional attention', or better its operationalisation in the relevant experimental settings, appears to be distinct from the one which underpins 'cognitive control functions.'[119] We should notice, however, that 'attention', in this literature, means shifts in perceptual attending, i.e., the so-called 'capture' of attention by some stimulus as due to factors related to the physical features of the latter, such as, e.g., 'colour', 'motion', 'contrast' etc. (cf. Pessoa 2013, 101). Insistence on exogenous characters of emotional attention in affective sciences is closely linked to the focus which characterises empirical research on how emotions influence the capture of attention by given stimuli–i.e., by facilitating or impeding it. In philosophy, I suggest, the insistence is due to the idea that emotions triggered *via* perception are paradigmatic. To wit, Brady (2014), who is one the most explicit proponents of the view that attention in emotion is distinctively exogenous, argues for the philosophical significance of the view as follows.

[119] See, e.g., Vuilleumier (2005) for an overview of the relevant literature. Cf. Faucher & Tappolet's (2002) and Brady's (2014) treatments of emotional attention.

> [T]here is significant epistemic value in an attentional shift that is passive, reflexive, non-voluntary, and effortless. One obvious value here is precisely that such shifts are effortless: if our attention is automatically and passively drawn to things that are relevant to our concerns, then we do not need to actively, continually, and consciously *scan the environment in order to detect such things*. Such voluntary attentiveness is typically very costly from the standpoint of our cognitive resources, and so it will be better, other things being equal, if attentional shifts were non-voluntary. […] It is therefore important for us to have ways of *registering* or *noticing* potentially important objects and events that are not (as) costly from the standpoint of cognitive resources. Moreover, the fact that such attentional shifts are automatic and reflexive suggests that there are advantages with respect to the speed of response, when compared with attentional shifts that are voluntary and active. Reflexive, automatic shifts of attention would seem to be quicker than conscious, voluntary, and effortful shifts, and there can be obvious practical advantages in *a fast response to potential danger and the like*. (Brady 2014, 19, italics added).

As I have mentioned, according to Brady, the essential function of emotion is a matter of 'alerting' or 'signalling': it is to *alert* us to the presence of concern-relevant objects or events in ways which allow us to dispense with *costly* information processing. In this sense, he is the polemical objective of the remark Frijda (2007, 27) makes about "the paradigm of emotion": he identifies it with "fear evoked by a sudden stimulus" as opposed to instances of interest, affective desiring, "lust" or "love".[120] In other words, although he is not, of course, exclusively interested in emotions that are sensorially triggered, Brady's driving intuitions primarily originate from such cases–just as the one that I owe to Frijda originates from cases in which emotional attention is deployed in cognitive contexts. In what follows, I will try to undermine the rationale for insisting on exogenous characters of emotional attention by suggesting that sustained attentional engagement with an item or topic is *always* endogenous,

[120] Cf. e.g., De Sousa (1987); Faucher & Tappolet (2002). Brady's view is discussed in more detail in chapter 6 in connection to the issue of how emoting subjects relate to value. As we will see, his treatment of emotional persistence is congruent with his view that the function of emotion is to allow us to detect: emotions continue so as to continue to alert us, thereby "promoting critical reflection about them" (op. cit., 92).

in cognitive and sensorial contexts, and that this does not prevent us from accommodating the exogenous character that theorists like Brady emphasise.

As I have mentioned in discussing the contrasts between the drive-based and sense-based kinds of affective reaction and emotional reactions, the distinction between endogenous and exogenous attention has been originally introduced by Posner (1980) as pertaining to the *locus* of control of attention-orienting: it regards whether such control is 'external'–i.e., due to the stimulus–or internal–i.e., due to the purpose of the individual. A nicety of this distinction as introduced by Posner is that it does not align the opposition 'external' *vs* 'internal' and oppositions concerning (in)activity or (in)voluntariness as is often done in psychology and philosophy of attention–cf. Allport (1993, 188) and Wu (2014, §1.7). The fact that Posner's distinction is conflated with distinctions concerning the relations between attention and the *will* explains, at least partly, why it is often taken to be essentially dichotomous. In my analysis of the structure of emotion episodes, I too have treated the endogenous *vs* exogenous distinction as a dichotomy. However, my argument was based on an important disanalogy between cognitive and sensorial attention: cognitive contexts do not afford a distinction between *pre-attentive* and attentive awareness, so that we lack the possibility of appealing to exogenous influences. But the endogenous *vs* exogenous distinction that one may draw within the limits of the sensorial contexts has different significance, and in this case the dichotomy is not as compelling. Roughly, the distinction is purely descriptive, for the capture of attention, plausibly, is always influenced to an extent one's current or standing purposes. The dichotomy is suggested by cases in which endogenous and exogenous control appear to be in conflict; for instance, cases in which purpose-*irrelevant* but salient stimuli distracts us, i.e., interrupt our sustained attending. But whether we take such cases to be cases of conflict between internal and external control will depend on our account of the relevant concept of 'salience', and more specifically, on whether we take such salience to solely depend on *intrinsic*

features of the stimulus as opposed to being partly dependent on the subject's (standing) purposes. In what follows I will present reasons for the second account of salience.[121]

First of all, we should notice that the notion of 'salience' finds its place in both cases: whether my attention is captured by a given sound because it is loud and sudden–i.e., exogenously– or because I'm listening out for it–i.e., endogenously–my attention is captured by the sound *because* the sound is *salient* to me. In the first scenario, the salience is apparently dependent on features of the stimulus, whereas in the second one, it is apparently dependent on my currently prevailing purpose. Let's call salience (apparently) dependent on features of the stimulus 'intrinsic salience' and salience (apparently) dependent on one's purpose 'relevance.' We can safely assume that, at any given time, any individual will harbour disparate standing purposes in the form, for example, of enduring interests, concerns, intentions etc. Whenever subjects are engaged in sustained attending, the activity is driven and governed by one of these various purposes–i.e., the one currently prevailing or taking precedence over the others. Let's imagine now that some stimulus, *irrelevant* to the subject's currently prevailing purpose, 'properly' distracts the subject, that is, it captures and then retains the subject's attention for a few instants before she can get back to her previous purpose. The capture of the subject's attention, in this case, is plausibly due to the intrinsic salience of the irrelevant stimulus. However, it is difficult to account for the ensuing *retention* of the subject's attention without mentioning some endogenous determinant: the stimulus, must have aroused an interest, however fleeting, for otherwise the interrupted attentional engagement would have been instantly resumed. In other words, the stimulus must have become relevant to some other

[121] The idea that the capture of attention by given stimuli is (at least partly) dependent on one's (standing) purposes is defended, among others, by Eilan (1998) and Roessler (1999), on whose discussions I rely heavily in what follows. They both spell out the way in which one's purposes bestow salience upon given objects in terms of *interrogative* attitudes. Cf. Evan's (1970) discussion of what he calls 'interrogative' attention. Koralus (2014) elaborates on the link between attention and 'interrogation' and puts forward as a general theory of attention as being essentially a "a matter of the relationship between questions (encoding the completion conditions for tasks for the purposes of monitoring) and what counts as answers to those questions." (47).

purpose. Roessler (1999) discusses this sort of view drawing on Neisser (1976)'s concept of a 'cycle' of perception. He defines the concept as follows:

> A perceptual cycle, in Neisser's sense, consists of three elements: an anticipatory schema, specifying what kind of information the perceiver aims, and expects, to acquire; exploratory activity guided by the schema, such as looking or listening; and finally the information picked up as a result of this activity, which in turn modifies the schema. (56).

Roessler's proposal is that, while it is true that our attention is sometimes captured in an exogenous–*stimulus*-driven–fashion, exogenous influences on attention-orienting are limited to the initial shifts: shifts *can* mark the inception of continued attending, but if they do it is because the subject has then engaged in an activity of 'looking at' or 'listening to' the stimulus, i.e., a sustained attending guided by a new schema. For example, I may come to *notice* the fly that entered my room because its constant buzzing distracts me from my writing–i.e., the fly may capture my attention; but if the fly then retains my attention, it must have become relevant to a new interest or purpose–e.g., I am now watching it fly around the room with the goal to lead it out of the window, or I am just observing how erratically it flies around.[122] A further weakening of the role of the exogenous influences on sustained attention is also defensible, because, arguably, the shift itself is never completely free from internal influences. The salience of all novel or abrupt stimuli, for example, may be explained by the fact that,

[122] A full defence of this view requires arguing that (sustained) attending to an item is attending to an item under an *aspect*, and that this selectivity is best explained by appealing to some 'schema' specifying what kind of information one *aims* and expects to extract from the item–viz., an interest in an aspect of the item. Neisser offers a clear illustration in the passage I quote below. "If we happen to see someone smile, for example, there may be information to specify (a) the shapes of his teeth; (b) the changing positions of his lips; (c) the fact that he is carrying out a certain culturally significant act; i.e., smiling; (d) something about his mood, which may be cheerful or sardonic or merely polite depending on the context in which the smile occurs. When we perceive his mood, we are not engaged in the same perceptual cycle as when we are attending to his lip movements. We develop a different (though perhaps overlapping) set of anticipations; we pick up information that extends over a different span of time […]." (Neisser 1976: 21-32, quoted in Roessler (1999, 56)).

even when such stimuli are irrelevant to currently prevailing purposes, they are still relevant to one or more of the standing purposes harboured by the subject–e.g., a standing interest in assessing the relevance of any novel stimulus (cf. Evans 1970, 86-87)[123]. On this view, what may at first seem clear cases of conflict between the subject's purposes and the intrinsic salience of given stimuli, turn out to be cases of conflict between two distinct purposes. The view that salience is always partly a matter of relevance to purposes is empirically supported by research on attention and performance in the tradition inaugurated by Folk.[124] Folk *et al.* (1992) challenged the widespread consensus that exogenous attention-orienting is free from endogenous influences, "at the mercy of the appearance of certain classes of stimuli (be they transients, abrupt-onset stimuli, new objects, or more generally, relatively salient stimuli)", which "draw attention in a completely bottom-up manner, overriding or at least delaying the cognitively driven, top-down control of attention" (Pashler *et al.* 2001, 633). Folk's key point of departure was the observation that studies of exogenous attention-orienting employed distractors instancing the same features instanced by the *targets* the subjects were supposed to be searching for. For instance, 'abrupt-onsets' distractors were used when the subject was supposed to look precisely for abrupt-onsets targets. This fact cast doubts on the view that the capture of the subjects' attention, in such experiments, was determined by the stimuli's intrinsic features, for it may instead be driven by the match between such features and the features used to find targets. To eliminate this potential confound, Folk in his own experiments controlled for the relation between features of the distractors and the features the subjects used to find their targets.

[123] "Because the environment is of *interest* and *concern*, every perceptible change in it taken by the percipient to be novel must alert it. […] Thus, the infant may be startled by a clap of thunder that an adult ignores. […T]he stimulus is still 'picked up […]. What the adult has inhibited is any further reaction–such as the assumption of a posture of attention.'" (op. cit., 86, italics added). Cf. Eilan (1998, 194-195).
[124] In the following paragraphs I rely on Pashler *at al.* (2001).

> They tested two unique distractor properties (color or abrupt onset) against each of two unique target-finding properties (again, color or abrupt onset). Confirming previous results, they found that when subjects are looking for abrupt onset targets, abrupt-onset distractors capture attention and color distractors do not. However, when subjects are looking for color targets (e.g. a red object in a background of green objects), abrupt-onset distractors do not capture attention, and color distractors (e.g. red distractors in a green background) do. (Pashler *et al.* 2001, 634)

Folk proposed that his results were best explained by what he named '*contingent* orienting', i.e., the hypothesis according to which the capture of attention is contingent on the match between the relevant features of the distractor and subject's attentional 'control settings'– viz., the subject's purpose.[125] Now, explanations in terms of intrinsic salience may still be the most apt when sustained attending is interrupted by the capture of attention by an irrelevant stimulus. In such cases we would normally describe the situation as one in which the subject was *distracted*, more or less momentarily. What constitutes distraction, however, will still depend on what specific purpose was taking precedence when the shift occurred; had the subject been engaged in activity aimed at detecting, say, all novel stimuli appearing in her visual field, the same shift would have qualified as continuation of sustained attending. In some cases, distractions turn out to mark off the inception of new episodes of attending; in such cases, it is apt to say the subject has acquired a new purpose. The new or newly activated purpose will then set its own constraints on what kinds of shift will constitute continuations *vs* interruption. The relations between the concepts of 'purpose', 'distraction', 'continuation' and 'interruption' will be examined later on in more detail. They are in fact essential to identify the criteria for the self-awareness distinctive of sustained attending. But this quick illustration of the idea that our purposes put constraints on what might constitute a

[125] See Pessoa (2013, ch.4) for more recent references to research in this tradition.

continuation of sustained attending suffices to suggest that the aptness of descriptions opposing intrinsic salience and relevance is precisely *descriptive* (Pessoa 2013, 101).

Let me now take stock and conclude this section by going back to the issue of the exogenous characters of emotional attention. If the point of insisting on such features, as Brady argues, is to ensure that we are endowed with affective systems which allow us to quickly detect concern-relevant objects and events when we are *not* looking for them, the general picture of the relations between purpose and salience I have sketched would seem to contain all the elements that are required to accommodate this passivity without denying the *interest*-driven character of sustained emotional engagement. The contingency of the capture of attention on one's 'control settings', be them set by deliberate purposes or affective interests, doesn't make the shift any less passive, reflexive, automatic, effortless etc.[126] On the other hand, ongoing stimulus selection of the sort that we observe in emotion seems to be the mark of all purposive behaviour, as opposed to purely reflexive behaviour: both kinds of behaviour are analysable in sequences of 'response to stimulus', but only the former results from being shut off from certain stimuli and more sensitive to other stimuli.[127] A system that may arouse in us affective interests, and then keep them 'alive' so as to allow us to engage in an active process of exploration of significance has arguably just as much epistemic value. Having made room for an account of sustained emotional engagement as *interest* driven, I will now turn to the general phenomenon of sustained attending. An upshot of the above discussion is the implicit suggestion that specifying the aspect under which mental agency is exercised in sustained attention requires articulating the interaction between the ongoing prioritisation

[126] Cf. Pessoa (2009; 2013, ch.4) for more references. Pessoa's work on emotion and cognitive control functions is, to my knowledge, the only notable exception to the current orthodoxy concerning the exogenous character of emotional attention. In his 2013 book, He argues that the "dichotomy exogenous/stimulus-driven/automatic *vs* endogenous/goal-driven/controlled, beaks down" and that such terms "can only be used descriptively" (102-103). His view is that while it is true that an *emotional* stimulus shall usually capture attention more easily and more robustly than a *neutral* stimulus, overriding the subject's explicit goals in given tasks, this does not mean that emotional attention is essentially *exogenous*; rather, it means that emoting subjects have emotional goals, which, typically, prevail over unemotional goals.

[127] This idea is not new. See Woodwhorth (1921, ch.4) for an early formulation of it in the context of his defence of an anti-behaviourist, motivational approach in psychological investigation.

of purposes *over time* and attention-orienting *at an instant*.[128] In what follows, I will articulate this implicit suggestion.

2 Attention and selection

In psychology and the cognitive sciences, the term 'attention' is used to identify a domain of research encompassing various attentional phenomena–e.g., task-switching, task-monitoring, alertness, orienting, visual-motor control, (dis)engagement etc.[129] Although there is deep disagreement about which single *sub-personal* structure underlies the phenomena associated to 'attention', it is widely agreed upon that they manifest two key features of our capacity for information processing: 'limitation' and 'selectivity'. The two features, as Pashler (1998, 2) puts it, "are really just the flip side of the same coin": since our mind can only concern itself with limited amounts of information at a time, it continually prioritises some of the available information. Although limitation and selectivity are theoretical constructs,[130] they are also part of our pre-theoretic conception of attention,[131] as is recognised by a number of theories of this phenomenon. Mole (2011, 4), for example, suggests that an account of attention is essentially "an attempt to explain the selectivity of our mental engagement with the world", Watzl (2011, 843) identifies attention as the "selective or contrastive aspect of the mind", and Jennings (2012, 536) states that "ordinary usage of 'attention' centres around a unique concept that is not picked out by any other term in English language: […] mental selection".

[128] Cf. Eilan (1998, 192): "one thing needed for getting the mix of activity and passivity right here [i.e., in perceptual attending] is getting right the relation between what the subject does when extracting answers to her questions from the environment, and the information-processing involved in her so doing. More specifically: the correction to the picture of perception as passive that is needed is one on which rational activity is said to have an influence on the processing of information in perception."
[129] Cf. Pashler (1999) for an overview of research on attention in cognitive sciences.
[130] They are part of the legacy of Broadbent's (1956) proposal that we apply to the study of attention the theoretical apparatus elaborated for telecommunication (Mole 2021, §1.6).
[131] Cf. Pashler (1989, 312-313), Peacocke (1998, 67-68), O'Shaugnessy (2000, 83).

Now, despite this widespread starting-point, attentional selectivity is theorised in many ways. Attention has been defined, to mention a few options, as the *process* continually structuring experience in foreground and background (Watzl 2011, 2017), as the *capacity* for–and *act* of–selection for action (Wu 2011, 2014), as the *state* of consciousness which makes information available for rational control of thought and activity (Smithies 2011), as an enhanced *sensitivity* in monitoring our progression in task execution (Koralus 2014), as 'unison' among various processes–i.e., a *mode* of occurrence (Mole 2011). Surveying this vast literature is well beyond the scope of this thesis. I am interested in examining one of the factors of the disagreement. Let me begin by setting aside factors which produce superficial disagreement. If we consider the above list of definitions, it may be striking, at first, that accounts of attention widely differ in their basic ontology of this phenomenon–namely, on whether it is a process, a capacity, an act, a state, a sensitivity, or a mode of occurrence–and in the (functional or phenomenal) characteristics they take to be essential to, or defining of the phenomenon. This ontological disagreement is at least partly explained by the fact we may approach attention under many of its *temporal* aspects. To illustrate, the act of directing one's attention onto an item marks the onset of an activity or process of attending to that item, and the activity or process is associated to a state of attention that will obtain for as long as the activity or process occurs, namely, the cognitive-epistemic condition the subject finds herself in for the duration the occurrence of such activities. That *states* of attention are described in terms of alertness, intentness, absorption etc.–all terms epistemically connotated–speaks in favour of this view. Furthermore, attentional selectivity is deployed in cognition, perception, and bodily action, and it may exhibit different functional or phenomenal properties in different contexts of deployment (cf. Mole 2011, 5). For example, perceptual attending, unlike cognitive attending, licences not only talk of engagement or occupation of the attention, but also talk of targets or objects of attention. But, as we have seen, this seems to depend on differences between

the sensorial and the cognitive and does not itself threaten the unity of the phenomenon.[132] The idea that attention is best seen as a 'manner' of occurrence of processes as opposed to a kind of process deserves further consideration. What remains to be clarified is the relation between attentional occurrences and conditions, on the one hand, and the concept of 'attentiveness'– i.e., of doing something attentively–on the other. The concept of 'attention' has a number of features that should incline us to deny that attending is a *sui generis* activity or process–an entity over and above the cognitive and sensorial activities that we are *attentively* engaged in. As White (1962; 1964) argues, attention is a 'polymorphous activity notion'. By this he means that although 'attending' typically signifies something that one does, it does not signify any 'doing', because what specific doing a subject's attending consists in is bound to what object the subject is attending to and in what context. In White's (1964) words, "to say that someone is attending […] gives us no more clue as to what activities he is engaged in than simply to say that he is practising. What 'attending' tells us is that his activities and energies, whatever they are, are directed to and focused on something which occupies him." (7, italics added). Accordingly, we may always replace the term 'attending' with terms referring to the specific activity the subject is engaged in.

> [B]ecause we attend to what is perceptible by using the appropriate sense-faculty and to what is intelligible by making it the object of our thinking, we find that phrases expressing the general notion of attention are interchangeable with phrases signifying these particular perceptual and intellectual activities. Thus we can "attend to" or "look at" the black-board, "attend to" or "listen to" the music; we ask someone to give us his attention or to lend us his ears, we catch the waiter's attention or his eye; we can give a good deal of thought or attention to a request and either our mind or our attention may be fixed on or engaged by a knotty problem. Therefore, when we speak of attention being paid or given,

[132] On reconciling the ontological variety with the unity of attention, see Watzl (2017, ch.2). On the contextual one, see, e.g., O'Shaughnessy (2000, 83); Pashler (1999, 312-313); Peacocke (1998, 67-68).

> drawn or attracted, it is basically […] these perceptual and intellectual activities to which we refer. (White 1962-63, 107).

Some (e.g., Mole 2011) have taken these facts to suggest that what 'attention' actually signifies is the 'manner-of-occurring' of activities or episodes or processes, and, hence, that we must opt for the so-called 'adverbial' account of attention. Mole's adverbial account is analysed in more detail later. For now, I just want to point out that an adverbialist treatment of attention is not necessitated by the claim that attention is a polymorphous activity notion. Attending is intellectual or perceptual, and when it is perceptual, it takes place in some sensory modality. However, that 'attending' and terms signifying intellectual or perceptual activities are interchangeable can also be interpreted the other way around: attentive activities are varieties of 'sustained attending', a mental category the unity of which consists in that its members all share the same kind of activity-*structure*.[133]

The factor of disagreement I'm interested in is one which stems from two ways in which instances of attention are identified on different accounts. On some accounts, instances of attention are identified in reference to the constancy of the presence to the mind of an 'object', whereas on other accounts, they are identified in reference to the constancy of the presence to the mind of a 'task'. These ways of identifying instances of attention, I suggest, tend to lead to privileging one of two ways in which the mind manifests selectivity, i.e., respectively, synchronically and diachronically. The synchronic selectivity of the mind is manifest at an instant, for at each instant only a part of the totality of the 'objects' of which we have pre-attentive awareness is at the centre or the foreground of our consciousness,

[133] Identifying this structure will be one of the aims of the next chapter. As we will see, the main problem with adverbial accounts is that they take the analysis of 'attention' to a level of abstraction that is too *high* for the structural differences between sustained attending and 'idle' ongoing attendings, such as, e.g., daydreaming to come into view. Roughly, this is due to the fact that the criterion for sustained attending that such accounts specify is a privative criterion–e.g., absence of task-irrelevant processing–and such criterion fails to distinguish sustained attending and idle ongoing attending, so that the distinction needs then to be secured in ways that are *ad hoc*.

while the rest lies in the periphery or the background (Evans 1970; Watzl 2011, 2017). Diachronic selectivity is manifest over time in the ongoing prioritisation of one of our purposes over other purposes. Some philosophers have suggested that these two ways in which the mind manifests selectivity are analysable in terms of each other, so that we can focus on either one of them and then extend our account to the other (cf., e.g., Watzl 2010, 114-116; Mole 2011, 24). I will reject this suggestion. The first reason to doubt that synchronic and diachronic selectivity are analysable in terms of each other is that accounts focussing primarily on either form of attentional selectivity attribute to attention mutually exclusive properties. According to the first family of accounts, attention is *ubiquitous*, that is, it is always present to varying degrees–viz., dispersed attention, as in mind-wandering or daydreaming, is an instance of attention. Relatedly, attention is taken to be an (in)voluntary phenomenon–i.e., it is taken to be, in principle, *subject to the will*, but not always willed. According to the second family of accounts, attention is an 'on-off' phenomenon–viz., instances of dispersed attention are cases of lack of attention. Relatedly, attention is taken to be a volitional phenomenon–i.e., it is a manner of willing. Why such features go together, that is, how the question whether attention is 'ubiquitous' *vs* 'on-off' and the issue of how attention is related to the will are related becomes apparent if we contrast the way in which attention is manifest in consciousness in the 'idle' *vs* the active subjects.[134] James describes the state of attention in the idle subject in his characterisation of what he calls a "confused, dazed, scatter-brained state which in French is called *distraction*, and *Zerstreutheit* in German" (1890, 404).

> The eyes are fixed on vacancy, the sounds of the world melt into confused unity, the attention is dispersed so that the whole body is felt, as it were, at once, and the foreground of consciousness is

[134] Cf Evans' (1970, §3.11) contrastive discussion of what he calls 'unordered' attention and 'interrogative' attention.

> filled, if by anything, by a sort of solemn sense of surrender to the empty passing of time. (ibid.).[135]

Although James says that in such a state our attention is dispersed, i.e., present in some form, he explicitly contends that such a state is "the real opposite" of attention. What does he mean by it? Evans (1970, 77) suggests that James is tacitly equating attention *tout court* and attention in its highest degree of focalisation or concentration. Since James holds the view that focalisation is the essence of attention, he thinks of the latter as an aspect of the mind that is only present sometimes–i.e., he thinks of it as an 'on-off' phenomenon–and the opposite of which is what we call 'dispersed' attention. A compatible, but different explanation of James' claim is the one I'm suggesting, namely that James takes such a state to be the real opposite of attention because he identifies attention with the state we are in when we are active and exercise agential control over the way in which experience unfolds over time. When we are idle and agential control is relinquished, we lapse into the state he describes. Over the next two sections I will examine in turn synchronic and diachronic selectivity and I will argue that an account of the way in which we exercise mental agency in attention crucially requires articulating the relationship between the two. In order prepare the ground for this account, to be presented in the next chapter, I will defend the following claims. Attention is a capacity for selection that operates both in the idle and the active subject–i.e., both 'spontaneously' and when exercised. When it is not exercised, that is, when we are idle, it is manifest in form of synchronic selectivity; when it is exercised, that is, when are active or engaged in sustained attending, it becomes also manifest in the 'thematic' and rationally intelligible continuity of diachronic selectivity. If we construe the concept of 'attention' as suggested, we can accommodate synchronic and diachronic selectivity in one single account at the appropriate level of abstraction.

135 James, The Principles of Psychology, Vol. 1, 403-404.

3 Synchronic selectivity

I will start with synchronic selectivity and will take as my case study Watzl (2010, 2011).[136] Watzl's theory is very well suited to my aims since, on the one hand, he explicitly contends that attention is an ongoing activity or process–the central phenomenon of attention, he takes it, is 'attending to something' understood as the temporally extended activity of holding attention on an object (2011, 43)–but, on the other, he individuates instances of attention *via* a criterion that applies primarily at times and only derivatively over time, i.e., in reference to objects of attention at an instant. Watzl's point of departure is the idea that attention is the contrastive aspect of the mind: attention is an activity, or process[137] structuring our consciousness so that some things are 'central' or in the 'foreground' while other things are 'peripheral' or in the 'background.' The idea that attention is contrastive is a way of spelling out its being involved in the selectivity of the mind; it is meant to capture the link between attention and selection *holistically*. According to Watzl, accounts of attention tend to miss the fact that attending to something also involves experiencing all that is not being selected in distinctive peripheral ways. Watzl's central proposal, Structuralim (S), is that "every form of attention consists in a structuring of one's conscious point of view. Conscious perceptual attention, conscious attention to one's bodily sensations, conscious attention to one's trains of thought should all be viewed as ways of structuring one's […] stream of consciousness." (2010, 126). Specifically, attention is the "process of creating and sustaining attentional phenomenal relations between the parts of your stream of consciousness", and, most notably, the relation of centrality *vs* peripherality (2011, 28). Although this and other

[136] Watzl (2017) offers a more complex account of attention in which he puts due emphasis on what I label diachronic selectivity. Arguably, however, an account of the relations between synchronic and diachronic selectivity is still missing. Roughly, the reason is that his new account of diachronic selectivity as it manifests in consciousness is still in terms of the relation of centrality-peripherality, which, as we will see, may only allow us to capture the *phenomenal* structuring which emerges from the exercise of mental agency in sustained attending, as opposed to *rational* structuring–cf. Watzl (op. cit. Part II, esp. ch.9) In this section, I will focus on Watzl's older work since I find it more manageable and for suits best my purpose..

[137] Watzl uses the two terms interchangeably. Since he maintains that attention (viz., structuring) is caused either voluntarily or involuntarily, activity and process are plausibly reserved, respectively, for the voluntary and the involuntary instances of attention.

formulations of S recognise the processive character of attention in that they qualify attention as a process, or activity and make reference to the flux, or the stream of experience, I will suggest that the structuring of the stream of experience captured by S is primarily synchronic and that more work needs to be done to do justice to the distinctive properties of the diachronic structuring of experience. The reason why that is so is that while Watzl recognises how the structure of what he calls the attentional space changes continually over time, S can only capture such changes *qua* changes in the relations between objects of *simultaneous* awareness: an object of which the subject has peripheral awareness at an instant may, at a later time, become an object of the subject's focal awareness. However, no account is offered of the connection between the successive structures. To see this, notice that Watzl defines the notion of an *object* of attention in terms of the relation of centrality *vs* peripherality:

> Some object, event or property in the world is the focus or object of your attention just in case you have an experience of it that is at the center of consciousness, where an experience is at the center of your consciousness during a certain period of time just in case during that period of time no part of your experience is central to it (or equivalently: just in case during that period of time it is not peripheral to any part of your experience (2010, 29-30).

So, although attention may be held focussed on one object for extended temporal intervals, as soon as it is directed onto another object, another instance of 'attending to something'– viz., a new structuring, takes place. This raises the question of how this account may capture the sort of continuity that episodes of sustained attentional engagement with an item or topic seem to possess despite more or less constant changes in the object of attention, e.g., exploratory perceptual searches, as opposed to perceptual monitoring of one single object, and any exploratory intellectual activity or complex reasoning, as opposed to, say, the active

production of unchanging mental images. Significantly, Watzl (2010, 268-301) only puts his theory to work in offering an account of what he calls the 'horizontal structures' of consciousness, structures which constitute the synchronic *unity* of our conscious experience. He does not address the issue of the nature of what we may call the 'longitudinal' or diachronic structure of experience. On the other hand, however, he does briefly develop an initially plausible adaptation of his account of attending to an object to diachronic selectivity. He calls such adaptation an account of 'attentive manners', that is, an account of what it is for one to do something attentively. Let me examine the account. Drawing on Peacocke (2000), Watzl introduces a distinction between proper 'objects' of attention and 'occupants' of attention. He defines the notion of an occupant of attention by delimiting its extension to activities performed attentively: "something occupies your attention *at a certain time* if and only if *at that time* this thing is an activity that you perform attentively." (115, italics added). Possible occupants of attention are thus limited to (mental) activity, such as, for example, wondering whether something is the case, calculating, imagining, trying to prove something, playing the violine and walking the tight rope. Having defined the concept of an occupant of attention in this way, Watzl proposes that one acts attentively during some period of time if, and only if, during that period of time, the activity in question, is a way of attending–that is, on his account, if it is a way of structuring the stream of consciousness. Watzl briefly illustrates this idea as follows:

> Consider concentrating your attention on a certain train of thought or reasoning (for example, concentrate your attention on calculating the following sum: 5444+7769). When you do that, you are structuring your stream of consciousness so that your awareness of a variety of mathematical relations will be at the center of your conscious experience (or possibly–if you are using visualization techniques–your awareness of a variety of material objects or events that you are imagining). Similarly, when you are attentively developing a logical proof for a certain proposition, your awareness of the logical relations among the various propositions will be at the center of your conscious experience. Attentively proving

> something (or inferring something) is different from non-attentively doing so, in that only in the former case will your awareness of the relevant logical relations be at the center of your conscious experience. (ibid., 114-115).

The idea is thus that attentively engaging in an activity involves structuring one's experience so that one's awareness of certain objects, namely *objects* of the *kind* required by the activity, is at the centre of consciousness at each moment for the duration of the activity. The first reason to think Watzl's account of attentive manners, as it stands, needs to be supplemented–if we are to capture the diachronic structuring of experience of attentive subjects–is that two clearly different activities may nonetheless involve the same kinds of object, and even the very same objects, which, on Watzl's account, entails identical structuring. For example, proving, studying, and rehearsing a given logical theorem are evidently different activities: they have different purposes, and they draw on different cognitive capacities. Nonetheless, they may well involve the very same objects: awareness of the very same logical relations might be at the centre of experience whether you are proving, studying for the first time or rehearsing, say, De Morgan's laws. Moreover, we may even build a case in which the same objects are at the centre of the experience of a subject who is idle or not engaged in any of the above activities. However unlikely it may be for one to absent-mindedly go through all the steps required by proving De Morgan's laws, it is not at all unconceivable that that might happen–for example, it may happen, perhaps, on the night before a logics exam while trying and failing to fall asleep after excessive rehearsing. This last case suggests that the main reason to be sceptical that Watzl's proposed account of attentive manners may capture attentiveness is that being attentive is not a matter of which (kinds of) objects may occupy the centre of consciousness relative to temporal intervals; rather, it is a matter of whether or not an aim

or purpose is causing consciousness to be structured around certain kinds of object(s).[138] Watzl's theory is silent on the relation between purposes and attention.[139] An account of attentive manners should indeed also allow for objects of utterly different kinds to feature at the centre of one's consciousness in different moments: pursuing a given thought, e.g., may involve recalling, inferring, supposing or visualizing a variety of 'objects', and the attention may come to be occupied by different objects and thought contents successively, such that if we didn't know *what* train of thought were being pursued, i.e., what the subject's aims were, we might as well take her to be absent-minded or else to be continually distracted while trying to do something. To sum up, Watzl's proposal that attention is the activity or process structuring our consciousness into centre and periphery is intuitively plausible and attractive if we focus *exclusively* on s-selection. Several authors implicitly or explicitly recognise that this sort of (phenomenal) structuring is an essential element of the phenomenology of attention: at any instant, some 'object' or other must occupy the centre of our conscious experience (cf. Evans 1970). But our consciousness also exhibits structure diachronically and this seems to be an altogether different sort of structuring that one should not expect to be able to capture by simply considering which objects happen to occupy the centre of consciousness. As we will see, diachronic structuring is not merely phenomenal; it is thematic and rational: it consists in the thematic continuity and the internal diachronic intelligibility of episodes of sustained attentional engagement. For this structuring to come

[138] Cf. Evans (1970, §4.2). Evan's account is an evident precursor of Watzl's in that he also defines attention as structuring of consciousness into centre and periphery. However, he supplements his definition discussing 'interrogative' attention as attention in which a given 'master-idea'–viz. a purpose– is structuring consciousness around certain objects.

[139] Peacocke (1998) makes this point arguing that mental agency has its distinctive contribution to make to subjective experience and that this contribution cannot in principle be captured by appealing to the objects or thought contents which successively occupy the attention of the active subject. "It is worth reflecting further on the striking fact about attention and consciousness that your attention can be occupied by your *trying to do something in thought*. Your state is subjectively different in the case in which you are trying, in thought, to achieve a particular kind of result from that in which you are casually drifting in thought. This can be so *even if the same sequence of thoughts occur to you in the each of the two case*s. Though striking, the point seems to be a special case of a more general phenomenon. In general, a subject's trying to do something (and what it is he is trying to do) contributes to what it is like for the subject. It does so in a way which goes beyond any occupation of attention by external events, sensations or thoughts." (op. cit., 68, italic added).

into view we must arguably individuate our unit of analysis in reference to something like tasks, aims or purposes pursued over time.

4 Diachronic selectivity

Turning to diachronic selectivity, I will take my cue from Mole's (2011) account. According to Mole, an account of attention should provide an answer the two following questions:

1. What is attention?

2. What is attentiveness?

1) is answered by specifying some kind of episode or process; 2) is answered by specifying some mode or manner of occurring of episodes or processes. Mole says that even though there are two questions to be answered, there is but one bit of philosophical work in need to be done, for once we have a clear answer to either one of the two questions, we shall find ourselves in position to immediately derive an answer to the other question. In Moles' terms, if we have an answer to 1), we may then derive an answer to 2) by developing the idea that "to do something attentively is to do the thing in the manner that typifies instances of the process of attention (a process we identified in our answer to question 1);" whereas, if we have an answer to question 2) then we may derive our answer to question 1) by developing the idea that "attention is the doing of something in the attentive manner (a manner we identified in our answer to question 2)" (op. cit., 24). Mole also claims, however, that even though an answer to either one of the two questions must immediately yield an answer to the other, we still need to make up our mind as to which of the two questions we should be prioritising, for 1) and 2) actually reflect two distinct *metaphysics* of attention, namely an 'occurrent' *vs* an 'adverbial' metaphysic.[140] In other words, they assign attention to two distinct

[140] Mole speaks of a process-first view. For consistency, I use my own terminology.

ontological categories: 'occurrent' and 'mode', or 'manner' of occurring. In a footnote, Mole considers the possibility that 1) and 2) may deserve to receive two independent or interdependent answers, but he quickly dismisses it as follows:

> Couldn't there be 'no priority' views, according to which the two questions get independent answers, or mutually dependent ones? I think not. Grammatical mastery alone seems to be enough to get one from an understanding of 'attention' to an understanding of 'attentively'. It is therefore implausible that the two questions could receive independent answers. The idea that the two questions are mutual dependent is harder to make sense of. It seems to imply that no noncircular analysis of attentional phenomena could be given, and so we may treat it as a position of last resort. (ibid., footnote 1, 176).

The reason I pause to notice this is that I will suggest we may also understand 1) and 2) as concerning, respectively, synchronic selectivity and diachronic selectivity, since to focus on the phenomenon of attentiveness is to prioritise instances of attention identified in reference to aims or tasks pursued over time. If we interpret Mole's contrast as opposing two ways of identifying instances of attention–*via* objects at times *vs* purposes over time–and not two different ontologies, the interdependence option quickly discarded by Mole actually becomes an appealing candidate. The idea here is that in order to account for attentional selectivity we need to articulate the relation between the two kinds of selectivity, for although they are different and not mutually reducible, they are interdependent: on the one hand, which given objects will be selected at given times partly depends on which purpose is being currently prioritised; on the other hand, the successful continuous prioritisation of one purpose at least partly depends on which kinds of object may *capture* one's attention as one pursues such purpose–i.e., on whether shifts of attention constitute continuations or interruption. Before I start to articulate this proposal, I will discuss Mole's account of attention and try to show that, although it has the merit of identifying instances of attention in reference to purposes,

and, hence, from a truly diachronic point of view, it is, like Watzl's, incomplete, in a way that does not allow us to locate the respect in which mental agency is exercised in attention.

Mole's starting observation is that some concepts, such as, e.g., 'haste', do not refer to any specific activity or process, but rather signify manners of occurring of activities or processes. He then goes on to argue that the concept of 'attention' is, in this regard, just like 'haste': describing someone's behaviour as attentive is not at all meant to advert to the involvement in such behaviour of any specific further process we may call 'attention', but to the *way* such behaviour is executed. My aim in what follows is not to assess Mole's particular proposal regarding the nature of the attention as an adverbial phenomenon, but only to draw some lessons from the adverbial approach as a kind of diachronic approach. Mole takes attention, or more precisely *attentiveness*, to consist in 'unison' among 'agent-level' cognitive processes, a state which obtains if, and only if none of the resources the agent would consider relevant to the task at hand, given their understanding of that task, is involved in task-irrelevant processing. In Mole's words:

> Let α be an agent, let τ be some task that the agent is performing, and call the set of cognitive resources that α can, with understanding, bring to bear in the service of τ, τ's 'background set'.
>
> α's performance of τ displays cognitive unison if and only if the resources in τ's background set are not occupied with activity that does not serve τ. (op. cit., 51).

Therefore, as Mole clarifies, "the condition that the cognitive unison theorist claims must hold over the background set if the subject is to count as attentive is a *negative* one: it is the *absence* of any irrelevant processing" (70, italics added).[141]

[141] The proposal is inspired by White (1964, 7): "Full attention to X consists not only in the range of activities that are centred on X but in the *absence of activities* concerned with things other than X. Hence, the common *privative* force of the notion of attention; one minds one's own business by not minding that of other people.

The main argument Mole offers in support of the adverbial approach, besides a detailed articulation of his 'unison' account and the exploration of its theoretical advantages, is an argument aimed to prove that an occurrent ontology of attention is ultimately untenable. The structure of the argument is as follows: any occurrent ontology must construe attention as some kind of ongoing process; it hence predicts no change from attention to inattention may ever occur without changes in the underlying attentional processes, that which makes the ontology vulnerable to compelling counterexamples. Having highlighted the limitations of an occurrent ontology, he then urges us to look at the prospects of an adverbial treatment. Now that I have outlined Mole's proposal and strategy, I can turn to examining his argument in some detail. The central upshot of the discussion will be that what Mole's argument actually shows is not that attention is adverbial, but that instance of attention ought to be identified in reference to purposes that are prioritised throughout temporal intervals.

Mole's recipe to generate his counterexamples is to suppose that attention is constituted by some kind of process (e.g., feature-binding processes, task-switching processes or any other), and then situate such process, first, in cognitive contexts in which, it *constitutes* an attending, then, in cognitive contexts in which, intuitively, it no longer does.

> We can imagine taking a set of cognitive events that constitute the paying of attention in one of attention's simpler instances; we can imagine moving those events, without changing any of the facts about which processes are taking place, and situating them in a new, more *complicated* cognitive context; and we can imagine this happening in such a way that, in the new context, the events are no longer attention constituting. The fact that such cases are imaginable is a problem for the process-first view. (op. cit., 36).

On the privative character of the concept of 'attention' see also Alexander (1953, 518).

Let me briefly consider one such case as discussed by Mole, namely the task of producing a string of random numbers. This task, Mole explains, proves extremely attention demanding, for the numbers which occur to us are not random; rather, the numbers tend to occur to us by following patterns. Consequently, "[t]o quickly generate random numbers one needs to […] switch between strategies of generation to prevent such patterns from developing." Therefore, he suggests, "when one pays attention to random-number generating, one's attention can be constituted by the processes involved in disengaging from one number-generating strategy and adopting another" (ibid., 39). Having specified the putative attention-*constituting* cognitive processes in this context, Mole then goes on to imagine a context in which the same kind of cognitive processes would no longer, intuitively, constitute attending.

> In a context where the subject's task is one that requires doggedly pursuing a single train of thought, the process of disengaging from that task and switching to another could, it seems, occur in ways that constitute failures of attention, not instances of it. If this is indeed possible then these processes […] constitute attention when the agent's task is random number generation, but *the same processes constitute inattention when embedded in a context where the agent's task is more cognitively elaborate.* (op. cit., 40, italics added).

Mole grants that such counterexamples are not conclusive, for the defender of an occurrent ontology of attention might question their cogence by appealing to some more finely-grained identification of the relevant processes or accepting that only in some cases given processes constitute attention, but not in all–i.e., by accepting that attention is not any single process, but a family of different processes. He insists, however, that as soon as we move away from the simplest cognitive contexts to consider more complex situations, the occurrent ontology runs into troubles such as this one and thus loses its initial appeal. His diagnosis of the issue with the occurrent ontology is as follows:

> The root of the problem that these particular cases point to lies in the fact that, in *sparse* and *undemanding* contexts, where the subject's task is simple and his time is short, the processing that constitutes his attention can be relatively simple. Once one of these simple contexts has provided us with an example of a simple attention-constituting process we can then construct a new example where we keep that simple piece of processing fixed but situate it in a more *complex* context, where the subject is engaged in something more *sustained* and *cognitively elaborate*. Where there is sufficient other processing to swamp the process that we started with […] we have an example where it is no longer feasible that that process constitutes the subject's attention. (ibid., 41, italics added).

Though I agree that Mole's cases reveal an issue with the way in which attention is identified, I do not think that the issue originates in the adoption of an occurrent ontology, and it is not clear what role the change from 'short and simple' to 'sustained and elaborate' tasks plays in the generation of Mole's putative counterexample, nor, indeed, whether Mole's putative counterexamples *are* more sustained and elaborate: the task of producing random sequences and the task of producing non-random sequences may take the same time, and Mole's criterion for judging of their cognitive complexity is not specified. We can begin to see where the problem lies looking more closely as what changes exactly from one context to another. In all the cases, what does the trick seems to be simply the fact that the subject's task has changed, and not that it has become more "sustained and cognitively elaborate." The subject now has a different purpose, and more specifically, a purpose that happens *not* to be served by the particular cognitive processes under exam. If the purpose is to randomly generate numbers, then, task-switching processes may constitute attention precisely because randomly generating numbers turns out to require a constant switching from one strategy to another, as Mole explains. However, the very same cognitive processes would constitute distractions if the task were instead to generate a sequence of numbers by means of one single strategy. The complexity of the task seems irrelevant; what really matters is a certain alignment

between the task at hand, on the one hand, and the processes employed to serve such task, on the other. In other words, whether the subject is paying attention or being attentive depends on whether what occupies her attention is in line with her purpose, and this does not seem to speak against an occurrent ontology; rather, it speaks against the possibility of assessing whether someone is attentive by means of identifying what engages their attention *independently* of their current purpose. Therefore, prioritizing question 2), rather than leading immediately to an adverbial treatment, leads, I suggest, to *identifying* instances of attention in reference to the purposes currently governing the subject's attention. Let me now turn to the question whether Mole's account can capture synchronic attentional selectivity. Mole maintains that one can in principle establish whether the pursuit of given tasks is 'object-involving' and which specific objects it involves by establishing whether the cognitive resources serving the pursuit of given tasks operate on representations of given objects.

> [T]o get an answer to the question of which objects the attentive agent attends to, our theory should start with attention to tasks, and should proceed from there to an account of attention to the objects that *participate* in those tasks. We can do this by defining 'object involving' as follows: Say that a task, τ, involves object o just in case the cognitive resources that serve τ are resources that operate on a representation of o. Most tasks will be object involving in this sense. (op. cit., 73)

He admits that taking attention to be primarily a mode of engagement with 'tasks' *vs* 'objects' makes it rather difficult to establish whether attentive subjects are attending to given objects in the pursuit of their tasks, and he says that, ultimately, in most contexts, there will be no way to determine whether an answer to the question of which object the subject is attending to is correct. He puts forward two options in the passage I quote below.

> A generous theory would say that, when o is an object involved in task τ, and when an agent is

> performing τ attentively, then and only then the agent is attending to o. A less generous theory would say that the agent is attending to the object, o, if and only if the task that is being performed attentively is one in which object o is, as we might say, *accusatively involved*, that is, involved as an object on which the task is being performed. The less generous theory would have the consequence that, when attentively hammering in a nail, the agent counts as attending to the nail but not to the hammer. The more generous theory would count both hammer and nail as attended. As far as I can see it is a matter of conversational pragmatics which of these standards is operative. There is no absolutely right answer about whether the hammer is among the things to which attention is paid when hammering in a nail. (ibid.).

While Mole is right to say that it is practically impossible to determine which specific objects occupy one's attention at each instant while one is engaged in a given activity, it still seems that once the task is specified we should be in a position to identify the specific constraints that such task puts on which object its pursuit might involve, for whether a task is being pursued attentively must partly depend on which objects occupy the subject's attention, and under which specific aspects, as well as on whether the subject's attention is directed on or captured by such objects *because* of the subject's attentive engagement in that task. Directing attention towards–or having it captured by–the shape or the colour of the nail, for example, is compatible with attentively hammering on the nail *if* it is caused, say, by an intention to check what material the nail is made of having realised that it is bending despite the fact that its inclination in respect to the wall is good. The same shift of attention, however, is not compatible with attentively hammering on the nail if it caused, say, by the thought that the nail might be one of the nails one used months earlier to hang one's favouring pictures when first decorating the apartment. In such a case having one's attention captured by the shape or the colour of the nail would count as being distracted from the task if only momentarily. The point is that it seems that we should be able to assess whether an instance of attending to an object, under some aspect or other, constitutes continued pursuit of the task at hand

or distraction. Simply saying that the nail is an object 'participating' in the task at hand doesn't discriminate between these two cases, but this seems something an account of synchronic attentional selectivity must be able to do. So, Mole's proposed extension of his account of attentiveness to attention to an object would seem incomplete, or at least too highly abstract. Now, Mole can concede, as he does, that his account is not fully equipped to account for attention to an object, while insisting that we do *not* need such an account, for we may still distinguish the two cases–i.e., continuation *vs* distraction–at the level of the relation between tasks at hand and processes: in the above example, thinking that the nail one is now using might be one of those one has used on another occasion constitutes task-*irrelevant* processing, and since per Mole's privative account of attentiveness, attentive φ-ing consists in absence of φ-ing-irrelevant activities, the presence of such thinking would count as distraction. In the next section, I will argue that we *do* need an account of attention involving an analysis of both diachronic and synchronic selectivity, as well as an analysis of the relation between the two, for the distinctive character of sustained attending as opposed to other forms of ongoing attending inheres exactly in this relation, that is, it inheres in in the way in which ongoing prioritisation of purposes determines what shall occupy one's attention at given times. My main argument will be that while emphasis on the negative or privative character of attentiveness secures the desired distinction between attentive and inattentive *task*-execution, it fails to yield a clear distinction between attentive task-execution and phenomena such as mind-wandering and daydreaming. An account *à la* Mole, which stays at the level of the relation between activities and cognitive resources, does not yield a distinction between sustained attending and *idle* ongoing attending. The reason is that by conducting the analysis at the level of the relation between the task at hand and task-(ir)relevant processes we cannot appreciate in which respects mental agency is exercised in sustained attending as opposed to the way in which it is exercised in idle ongoing attending.

5 Idle ongoing attending and spontaneity

Let me first notice that one sense in which Mole's account would be considered incomplete by many theorists working on attention is that, *pace* Mole, who contends that attention is an on-off mental phenomenon, such that when one is not attentively executing any task, attention is not operative in any way, it seems quite impossible to avoid appealing to attention in describing what is going on when an agent gets distracted, whether or not the distraction is momentary. The most natural way to describe such cases is to say that the agent's *attention* is drawn to or captured by something irrelevant to the task at hand. Now, such cases would only fall within the purview of Mole's account if whatever captured the agent's attention then became an object involved in or participating in some new tasks, say, a perceptual exploration of it. Mole is explicit on this point:

> One's attention can be caught involuntarily by a sudden sound at the window. If the way in which that episode of looking to the window unfolds is guided by one's understanding of what it is that one is doing in looking to the window, then the looking is a task, although not a task that one performs voluntarily. Because it is a task the looking is something that can be done attentively. If, on the other hand, one is distracted by a sound at the window without the processing of that sound being integrated into any understanding-guided activity, then the case is one where the sound is merely distracting and is not itself attended. (Mole 2011, 53, italics added).

But this restriction is quite artificial, and theory driven; it stems from Mole's view that attention, as an adverbial phenomenon, that is, a manner of occurring of activities, is exclusively connected to task-execution. If unpersuaded by Mole's ontology of attention, one may well grant that a momentary distraction involves transient attention to the distractor. Relatedly, many theorists working on attention wouldn't be satisfied with Mole's treatment of phenomena which intuitively constitute *idle* ongoing attending. Mole delimits the concept

of 'task-execution' in a way that is meant to exclude phenomena such as mind-wandering and daydreaming from the purview of his account, but wandering and dispersed attention still exhibit synchronic selectivity, which for many is best accounted for in terms of the contrast between what currently occupies attention and what constitutes our pre-attentive phenomenal background. In so far as the particular attentional phenomena I'm interested in are episodes of sustained attentional engagement with an item or topic, i.e., purposive, sustained attending, as opposed to transient attention and *idle* ongoing attending, the purview of the account I aim to put forward is as restricted as that of Mole's account, so that this is not why I take Mole's account is as it stands to be ill-suited to my purpose: I may reject Mole's theory-driven restriction of the purview of an account of attention and still accept Mole's 'privative' definition of attentive task-execution as defining purposive sustained attending: purposive sustained attending consists in absence of task- or purpose-irrelevant processes. The issue with Mole's account is that while a negative or privative condition for attentiveness is a necessary condition, it is not sufficient: it doesn't actually exclude idle ongoing attending, or so I shall argue. The feature distinguishing purposive sustained attending from other forms of attending inheres in fact in the *way* the subject controls attentional selectivity, i.e., the precise respects in which they are *active*, and, relatedly, the sort of self-knowledge associated to this species of mental activeness. Since it rests on denying that attention is operative in any way as soon as we move beyond the domain of task-execution, Mole's account is necessarily insensitive to differences in the way in which attentional selectivity is controlled by the subject when such control 'enforces' constraints set by some global reigning purpose as opposed to when it does not. To capture such differences, we need to go down to the level of the relation between d-selection and s-selection, as I will now try to argue. Mole does acknowledge that phenomena such as mind-wandering and daydreaming represent potential problems for his account. The main difficulty, according to him is the following:

> All sorts of processing resources could serve the subject's daydreaming, and all sorts of activities are such that, if they were to occupy those processing resources, then they would contribute to the subject's daydreaming. More or less any state of the daydreamer therefore looks like it will answer to our definition of 'unison'. This raises the threat that no theory of attention based on cognitive unison will be able to classify cases of inattentive daydreaming as being inattentive. (op. cit., 57)

His reply to this threat is simply to exclude phenomena such as daydreaming from the remit of his account by claiming that they do not normally constitute 'tasks', and hence are simply *not* susceptible of occurring attentively. His argument presupposes the above characterisation of phenomena like daydreaming as distinctively *unconstrained*–that is, such that all sorts of processing resources could serve them. The idea is that since daydreaming is unconstrained it cannot in principle be guided by the subjects' understanding of 'daydreaming', but this sort of guidance is an essential requirement for attentive task-execution.

> The threat here is merely apparent because the cases in which there really are no constraints on which processes do and which do not serve the daydreaming are cases in which the daydreaming is not under the control of the subject's understanding. *In such a case the subject does not have daydreaming as a task at all.* If so, then daydreaming is not something that this subject can do with cognitive unison, and there is no risk that a cognitive-unison–based theory will count the daydreamer as performing attentively. (ibid. italics added)

The threat, however, seems more serious than Mole recognises. The concept 'daydreaming' *does* put constraints on which processes may and which processes may not serve daydreaming. For example, if one is daydreaming, noticing a burnt smell coming from one's kitchen, realising that one might be late for class, or realising that one has walked past one's intended

destination would all qualify as *daydream-irrelevant* processing. If daydreaming is constrained in this sense–i.e., as the *kind* of activity it is–then, it could, in principle to an extent be guided by an understanding of what constitutes daydreaming. The fact that one may, for example, wilfully ignore the burnt smell coming from one's kitchen–assuming perhaps that someone else will take care of that–and thereby shut down the daydream-irrelevant processing, so as to continue to enjoy one's daydreaming, suggests that daydreaming might in some cases be guided by some kind of understanding. Now, Mole would reject this sort of counterexample. While he should concede that his characterisation of daydreaming is incorrect, for it is just not the case that "*all* sorts of processing resources could serve the subject's daydreaming, and *all* sorts of activities are such that, if they were to occupy those processing resources, then they would contribute to the subject's daydreaming", he may still deny that daydreaming is under the control of the subject's understanding of such constraints *unless* one has it as one's task. Mole does mention that one may daydream as part of a "deliberate relaxation technique, or with the intention of opening [oneself] to inspiration," and clarifies that he is only concerned with cases in which one just lets one's mind wander "in an idle moment" (ibid.). It is not clear, however, why the two cases would differ in any relevant respect: whether or not one has daydreaming as one's 'task', as O'Shaughnessy (2000, 218) observes, "one can scarcely so much as intend to daydream, mostly one slips into the state and discovers oneself adrift in the process a little later", therefore, in any case, "the process cannot be intentional under 'day-dream"'. In other words, although one may have daydreaming as one's 'task'–in the sense that one may *wish* or *desire* to slip into such state–daydreaming itself would remain essentially non-purposive. If daydreams are problematic for Mole's account, they are so in virtue of the kind of activity they are, not only in case the agent *wants* them to occur. That daydreaming has its constraints, such that one could control one's daydreaming with understanding, at least in principle and to some extent, I suggest, suffices to show that the phenomenon does represent a real difficulty for Mole's account.

Moreover, as we will see in more detail in the next chapter, although daydreaming activities are never intentional under the heading 'day-dreaming', they may be intentional under other, more 'local' descriptions in a way that seems to entail some guidance with understanding, which is all that Mole requires for an activity to be attentive. According to O'Shaughnessy's characterisation of daydreaming,

> At each point behind the scene lies an *intention* and its expression *in the will*. That causally active intention is at that very moment *determining* the next stage of the process, and so playing an *explanatory* role in relation to the advance of that process. And that is to say, that the subject himself is 'behind' it all. (op. cit., 224, italics added).

Soteriou (2013), in commenting the above passage, suggests that "[t]hought of in this way, we might say that in the case of daydreaming, one imposes constraints on ensuing mental activity in a relatively unconstrained manner" (326), and that "if O'Shaughnessy is right in his characterization of this activity, there is nonetheless a degree of practical self-knowledge" (ibid.). If this is correct, differentiating purposive sustained attending from daydreaming takes more than Mole's privative condition on attention. One might suggest what it takes is simply emphasising that daydreaming, unlike purposive sustained attending is 'spontaneous' in the sense of being unanticipated in the way it unfolds; although one may wilfully indulge in daydreaming, and to that extent control its continuation with understanding, one cannot control how it will unfold. If one did, one's activity would cease to be daydreaming. This suggestion seems to be on the right track, but it needs to be elaborated, and Mole's account does not have the resources to do that: to capture the spontaneity of daydreaming we must go *down* to the level of the relation between d-selection and s-selection. The reason is that purposive attending is also spontaneous in this sense. As Peacocke (1998) observes, if we are engaged in active, 'directed' thinking–that is, an instance of purposive sustained attending–

the thought contents which come to occupy our attention *occur* to us, that is, they are passive happenings. In fact, there can be no intention in advance to entertain *particular* thought contents, but only "intention to think a thought which stands in a certain relation to other thoughts or contents. It is thought carried out in accordance with such an intention that is directed, as opposed to idle." (68). O'Shaughnessy (2000) makes the same point by clarifying that to say that in active thinking one takes "active charge" of one's thoughts is obviously not to say that one "actively determines their content, which would be at once omnipotent, barren, self-refuting, and logically impossible"; rather, "[one] selects the direction of their movement, which is to say the content of the governing enterprise." (89). To insist on this point is to insist that even the sort of mental activeness involved in purposive sustained attending leaves room for the "unceasing spontaneity of thought, the continuing unanticipatedness of one's next thought." (201). Simply saying that idle ongoing attendings are spontaneous will thus not help us to mark the distinction. Mole's account is silent on this feature of sustained attending, namely, on the fact that being engaged in attentive activities over time is not just fully compatible with, but *essentially* involves passivity on the part of the subject as regard to the *upshot* of such activities. But sustained attending differs from phenomena like 'daydreaming' in the way in which the active ongoing prioritisation of purposes over time determines what shall tend to *occupy* our attention at given times–i.e., which objects, under which aspects, we may notice, and which thoughts may occur to us if our engagement is to be preserved. This seems enough to motivate us to look for another account, one which individuates instances of attention in reference to purposes over time, like Mole's does, but goes beyond Mole's account in clarifying the respects in which purposive, sustained attending differs from idle ongoing attending.

Tu sum up, Mole's arguments in support of his adverbial treatment of attention are best viewed as showing that attentiveness is a matter of the right sort of match or alignment between the task at hand and the cognitive processes deployed in its service. This does not

speak in itself against an occurrent ontology of attention; rather, it speaks against the possibility of assessing whether or not someone is being attentive independently of an assessment of their current purposes. Consequently, it speaks against the possibility of assessing whether someone is being attentive *at an instant* independently of an assessment of the relations between what occupies their attention at that instant and the continuous occupation of their attention over extended temporal intervals. Mental agency in attention is essentially exercised over time. We may see more clearly what the relation is between identifying instances of attention is reference to purposes and the diachronic perspective by considering the following example. Let's suppose that my attention is captured by some shiny black object partially hidden by the grass at time *t*. My attention to that object at *t* is in principle perfectly compatible with *both* attentiveness and inattentiveness at *t*. Whether or not I am attentive at that time depends in fact on whether my attention is captured by that object *because* of a purpose I am pursuing. If I am looking for the phone I lost in the grass–a shiny black object–my attention to a shiny black object at *t* is accountable a continuation of my visual search, and, possibly, if that object turns out to be the phone I've just lost, the completion of my visual search. But, if the object in question captures my attention while I am talking to my friend, listening to his travel stories, for example, my attending to the same object would count instead as a distraction. Now, the two cases are not distinguishable if our unit of analysis is an instantaneous occurrence, a temporal *snapshot* of my sustained attending to an object, for my attending to the shiny black object at *t* does not bear any trace of its causation, that is, it does not bear any trace of whether it is caused by my purpose as opposed to salience the object enjoys in virtue of some other factor. At this point, it may be tempting to interpret the contrast between sustained attention and any other form of ongoing attending in terms of the contrast between the 'agential' and the 'passive': sustained attending is actively guided whereas phenomena like daydreaming are instances of 'unguided attention' (e.g., Irving 2016). Our mental agency, however, is also arguably exercised in idle attending;

the difference, I will suggest, lies elsewhere: not in the *mere* exercise of mental agency but in the specific manner in which agency is exercised–and the state of self-awareness associated these different sorts of mental activeness.

Chapter 5

SUSTAINED ATTENTIONAL ENGAGEMENT

Introduction

The aim of this chapter is twofold: to specify criteria for sustained attention as opposed to idle ongoing attention, and, secondly, confirm and further elucidate, in light of such criteria, my central hypothesis–i.e., that sustained emotional engagement is a species of the former. Drawing on O'Shaughnessy's and Soteriou's proposals on where to locate mental agency in thinking, I will suggest that the mark of sustained attending is the imposition of constraints which make room for the applicability of the concept of 'distraction', and, relatedly, the obtaining of a state of self-awareness manifesting a sense of which transitions may constitute continuations *vs* interruptions of sustained attentional engagement with an item or topic. Sustained emotional engagement meets both criteria, or so I will argue. I will then identify two sub-classes of sustained attending to which sustained emotional engagement appears to belong. As I have anticipated, the first of the two sub-classes is characterised by 'exploratory *open-endedness*', whereas the second one is characterised by 'non-instrumentality'. In the final section of the chapter, I will elucidate the concept of non-instrumentality in reference to the contrast between *will*-driven and absorbed attending and I will relate the latter to various notions of 'interest' so as so clarify in what sense sustained emotional engagement is driven by *interest*.

As I have mentioned, O'Shaughnessy maintains that mental agency is present across all

'wakeful' and cognitive experiential continuities, including not only the ones which constitute purposive attendings, such as what he calls 'exploratory ratiocination', but also what he calls 'active idle drifts', namely a subclass of agential cognitive processes comprising borderline phenomena like soliloquy, daydreaming and hypnagogic mental activity (O'Shaughnessy 2000, 2016-2019). The strategy he employs to establish this view has two steps: first, showing that such kinds of experiential continuity all exhibit "a *degree* of practical self-knowledge" (Soteriou 2013, 326), and more specifically, what he calls "pellucidity", i.e., *accessibility* of the rationale which governs their progression over time, and second, showing that this sort of internal intelligibility is essentially dependent on an ongoing exercise of mental agency. O'Shaugnessy's core idea here is that mental agency "[i]imports internal intelligibility into the processive advance of the stream of consciousness […] through enabling reason to determine that advance (op. cit., 226).

> It seems that the only experiential processes that are rational progressions are active in character, and that *only through rationality, and so through rational agency, can internal intelligibility enter experiential processes in the self-conscious*. Only a mind steering its own cognitive path through a wider cognitive scene, a self-causing which is furthered by rational steps, can *introduce pellucidity into the flow of experience*. (op. cit., 226-227).

My aim in what follows will not be to assess O'Shaugnessy's claims about the 'range' of mental agency–i.e., that mental agency is involved in *all* wakeful cognitive processes,[142] including cognition that most philosophers and cognitive scientists classify as passive.[143] Rather, I will draw upon O'Shaugnessy's insightful descriptions of wakeful cognitive processes which seem to involve mental agency to distinct degrees–i.e., 'daydreaming' and 'exploratory ratiocination'–with the aim of extracting categorial *criteria* based on which one

[142] This is done in much detail by Soteriou (2013, Part II).
[143] For full defence of this view, Irving (2016) Irving & Thompson (2018); Cf. Irving & Glasser, eds. (2020).

may differentiate sustained attending and idle ongoing attending.[144]

1 Sustained attending and self-awareness

I will follow the structure of O'Shaughnessy's discussion of 'wakeful cognitive processes' and start with daydreaming. O'Shaughnessy argues that we do exercise a certain degree of agential control over daydreams' progression over time by suggesting that such phenomena bear the *mark* of the presence of–or guidance by–governing intentions, namely, a degree of *foreknowledge* on the part of the daydreamer concerning the progression of their daydreaming. Let me illustrate. As I have mentioned in the preceding discussion of Mole's treatment of daydreaming, O'Shaughnessy maintains that one cannot intend to daydream; one may obviously want or wish to daydream and create conditions ripe for daydreaming, but the ensuing mental activity will not be intentional under such description. Moreover, daydreams are essentially unpredictable in their course of development so that they cannot be plausibly considered intentional under any *global* definite description. For example, one may start one's daydreaming by remembering one's travels to some foreign country and end up fantasising about how one's life might then look like if one had decided not to come back. Nonetheless, O'Shaughnessy claims, "[a]t each point behind the scene lies an intention and its expression in the will", and this may be captured by allowing daydreams to be intentional, so to say, under more 'local' descriptions, that is, under descriptions which may apply to stretches of their overall course over time–e.g., 'recollecting the time when …', 'imagining how things might have gone if …' followed by 'anticipating the time when …' and so on. The presence of these intentions entails foreknowledge on the part of the daydreamer of the progression of their daydreams to the extent that *fleeting* intentions may still determine at each moment "the next stage of the process [and so play] an explanatory role in relation to the advance of

[144] The discussion in the next section owes much to Soteriou's (2013, ch.14) discussion of the place of agency in thinking and his presentation of O'Shaughnessy's ideas on this issue.

that process" (224). Let's assume that O'Shaughnessy's description of daydreams is correct, that is, that daydreams are governed by intentions at least *locally*, or moment by moment.[145] The distinctive characters of daydreaming as opposed more central cases of mental agency, O'Shaughnessy suggests, stems from how such governing intentions relate to one another and, crucially, the mode of production of the transitions: in daydreaming, "each intention is continually giving way to a successor intention, yet without their expressing themselves in discrete stages, and without their realizing any single *comprehensive* intention as they do. […]. This is because the connective tissue of these rapidly changing intentions is mere *association* and *inclination*." (217, italics added). In this characterisation, I suggest, one may find all the required elements to introduce a categorial distinction between idle ongoing attending and sustained attending. O'Shaughnessy is more interested in the analogies between idles drifts and purposive attendings than he is in insisting on their specific differences, and he suggests that idle drifts, from certain cases of 'soliloquy' to 'daydreaming' and 'pre-sleep phantasies', find their place along some continuum of degrees of agency, and, relatedly, self-knowledge. One extreme of the continuum is fixed by what I have been calling sustained attending, whereas the other one is the inner limit of wakeful consciousness as marked by hypnagogic mental activity or 'pre-sleep phantasy'–the outer limit being the dream. But the idea that what we have here is a continuum of degrees of agency is compatible with the suggestion that we can identify criteria differentiating idle ongoing attending and sustained attending. The key idea that one may exploit to do so is the idea that the *local* intentions in daydreaming do not realise any global, higher-level intention or purpose; therefore, from the agent's point of view, they are not assessable as continuation *vs.* interruption of daydreaming. As I have suggested, daydreams are susceptible of being wilfully continued, but the criterion for 'distraction' or interruption is set by the general category 'daydreaming' and not by anything inherent to

[145] See Soteriou (2013, 326-327) for discussion.

particular instances of daydreaming. In other words, even though there's some scope for the notion of being distracted *from* one's daydreaming–e.g., if a writer considered daydreaming beneficial to their creativity, she may try not get distracted by any task that may come to mind in her daydreaming–there is no scope, so it seems, for the concept of being distracted *within* one's daydreaming: no shift in governing intentions, and no ensuing change in the occupation of attention would be taken by the daydreamer to constitute a *distraction*. What underlies this intuition is the fact that daydreaming is incompatible with controlling one's attention in accordance with some *overarching* purpose. Even if one starts to daydream as a result of having set for oneself 'daydream' as an objective, the transitions in intentions are always caused by association or inclination. If they were caused instead by recognition of their conduciveness to the attainment of one's objective, one's mental activity would cease to be of the kind 'daydreaming.' By contrast, if one is engaged in purposive sustained attending, all transitions in governing intentions are governed by a global higher-level purpose, and the concept of 'distraction', accordingly, applies not only in cases in which one, say, slips into daydreaming, but also in cases in which one shifts from one purpose to another. For example, if my high-level purpose is to outline the bare structure of an argument from its premises to its conclusion, trying to come up with some cleaver examples illustrating one of its steps instead of moving onto the next step, would count as distraction, and this just as much as, say, looking up some references on Google Scholar or trying to chase a bee out of my window. Had the purpose been another, for example, getting a clearer grasp of an argument, including its structural elements and forcefulness, trying to come up with examples would have been accounted a continuation of it–but not so, say, looking up references on Google Scholar.[146]

Crucially for my purposes, the idea that sustained attending, as opposed to idle attending,

[146] Cf. Koralus's (2014, 45-47) for an analogous treatment of the relation between 'tasks' and 'distraction' focussing on the epistemic significance of the phenomenon.

involves imposed constraints on what might count as continuation *vs* interruption is not just meant to specify a third-person criterion for this phenomenon; it is meant to capture a feature of the state of self-awareness associated to it: the subject has a *sense* of what might qualify as continuation *vs* interruption of the engagement. One last clarification in order is that this sense reflects constraints which might be more or less stringent depending on the relevant purpose. It's important to take note of this because although specificity is a matter of degree, we may identify a threshold separating those purposes which specify completion conditions and those which, instead, only specify a certain general 'direction' of thinking. The existence of the threshold is reflected by differences in self-awareness. If one's purpose is an achievement, one will take it to specify completion conditions; relatedly, one will take one's engagement to be assessable in terms of failure or success, and one will monitor it accordingly–namely, to assess if the purpose has been achieved or to what extent it has been approximated.[147] If one's purpose is instead exploratory, as is the case in much of the thinking involved in doing philosophy, for example, the imposed constraints are more elusive, but they are nonetheless manifested in the subject's *sense* of what might constitute continuation *vs* interruption of engagement. As we will see in the next section, sustained emotional engagement is arguably an instance of exploratory sustained attention.

Now that we have obtained the desired criteria for purposive sustained attention–i.e., criteria which, unlike Mole's purely privative criterion, allow us to distinguish it from daydreaming–I may conclude my analysis of the general structure of sustained attention by going back to the relationship between diachronic selectivity and synchronic selectivity.

I will start by considering the following passage in which O'Shaughnessy characterises *exploratory* thinking articulating the interplay between the agential control we exercise in

[147] See Roessler (1999, 57-61) for discussion of this kind of purposive attending and the concept of 'monitoring.' Cf. Evans (1970, 104) on the distinction between 'unordered attention' and attention to which 'success-fail' conditions seem to apply.

sustained attending and the spontaneity of the outcomes of the ongoing process.

> Ratiocinative exploratory thought shares this much with fishing, that one is in either case dependent upon whether anything of value lies within a given region and on whether any such item is accessible. In each case one engages in an activity in the hope that a desired entity 'surface' to cap one's efforts. […] Just as listening to a speech prepares a mental matrix for the reception of intelligible sentences relating intelligibly to those preceding it, without either anticipating the sentences in advance or being wide enough to accommodate the noise of fireworks […]; so exploratory thinking imposes restrictions such that some astonishing mental occurrences will be accounted continuations of the ratiocinative process, whereas some others–such as a stray memory of a cricket match–will not. […] *In neither case is explanation in terms of mental purposes deflected.* Even though unknown resources of the mind play an essential part in throwing up into consciousness the relevant thought θ as one puzzled over xyz, and despite the fact that the functioning of those resources is not one's thinking, it remains an unqualified truth that thought θ entered the mind because one was thinking about xyz. (O'Shaughnessy 2000, 221, italics added).[148]

O'Shaughnessy's description confirms the intuition articulated at the end of the previous chapter, namely that to know whether a subject is (in)attentive at a time *t*, we should ask *why* the subject's attention is occupied by whatever it is that occupies it at *t*. If an answer in terms of purposes is available to the subject, such as, for example, "I thought such thought at *t* because I was puzzling over xyz", we may then infer that she was exercising agential control over her capacity for attention; if no such answer is available, for example, if the subject's answer is "I don't know" or "O was bright and sudden" or something like "an association of ideas made me notice O while I was actually puzzling over xyz"–we may conclude the subject's capacity for attention at *t* was idly operating. In general, whether or not the subject

[148] Cf Peacocke's (1998) discussion of 'directed thinking.'

is attentive at *t*–i.e., whether she is exercising agential control over her capacity for attention at that time–is only assessable if we consider an extended temporal interval (I) comprising *t* and then ask whether the occupation of the subject's attention at *t* is explained by some global purpose governing the subject's attending throughout (I). In the previous chapter, I began my discussion of synchronic and diachronic selectivity by observing that accounts of attention which focus on either form of selectivity at the expense of the other tend to ascribe to attention opposite features. In synchronic-selection theories, attention is *ubiquitous* although it is present in varying degrees and it is *subject to the will*, whereas, according to diachronic-selection theories, attention is 'on-off', and, relatedly, it is *volitional*–i.e., a manner of willing. We can now see how one may reconcile the two ways in which the mind exhibits selectivity within one single account attention. As anticipated, the key to doing so is to understand attention as a capacity for selection operating both in the idle and the active subject and over which agential control is exercised *essentially* over time.[149] If no control is exercised, as it is the case in mind-wandering or daydreaming, diachronic selection is inoperative (or perhaps it is operative at best locally, moment by moment). In such cases, the selectivity of the mind is (primarily) manifest as synchronic selectivity. If we do exercise agential control as in sustained attending, the selectivity of the mind becomes also manifest diachronically, that is, it is reflected by the ongoing thematic continuity and intelligible diachronic connectedness of successive instances of synchronic selectivity as belonging to one instance of diachronic selectivity, that is, an episode of sustained attention. In such cases, the progression of the stream of experience is constrained throughout, for all instances of synchronic selectivity that succeed one another do so in accordance with some *high-level* purpose. In the next section I will go back to sustained emotional engagement to reassess my working hypothesis–i.e., that it is inherently purposive–in light of the criteria specified.

[149] I borrow this way to put the point I'm trying to make from Evans (1985) on 'keeping track of something'– "[…] one's thought *at* a time is dependent upon an ability which is necessarily manifested only *over* time." (310).

2 Sustained emotional engagement

The idea that sustained emotional engagement is inherently purposive was initially suggested by elaborating on Frijda's hypothesis regarding the function of emotion, namely that this function is to cause action tendencies aimed at coping or dealing with emotional situations and ensure that they enjoy control precedence. In chapter 1 of this thesis, I suggested that sustained emotional engagement is in fact naturally understood as the mental aspect of emotional action readiness: it provides for the diachronic selectivity plausibly required to protect the actions from possible interferences and ensure the attainment of the emoting subjects' relational purpose–i.e., the purpose to maintain, establish, or disrupt a certain relationship with the environment. If persuaded by this general picture of emotions as control *systems* geared toward action it is natural to suppose that emotional cognition is inherently purposive, as opposed to being governed by ideational association and inclination like mind-wandering or daydreaming are.[150] In the previous sections, sustained attending was shown to involve an active prioritisation of some global purpose resulting in the imposition of determinate constraints on the occupation of attentional resources, constraints that are operative, not only locally or moment by moments, but also globally, that is, at the level of the transition between the sub-intentions the subject may acquire in pursuit of her purpose. The mark of this variety of mental activeness was identified with the obtaining of a state of self-awareness which makes room for the sense of what constitutes a distraction. Now, sustained emotional engagement, I suggest, is similar in all regards: the attitudinal-evaluative components of emotion episodes– viz., our emotion attitudes as variants of interests–are fit to play the role of sub-intentions, and the changes in attitude which occur in multi-attitude emotion episodes are naturally thought of as results of an ongoing process of exploration of

[150] It is perhaps interesting to notice at this point that O'Shaugnessy (2000), in his discussion of daydreaming, remarks in passing on the relation between daydreaming and emotion in a way which suggests incompatibility. "A certain frame of mind seems to be needed if [daydreaming] is to occur. It is difficult to dream with anxiety." (218).

significance as opposed to changes due to our associations and inclinations. Let me illustrate. During the course of an episode of 'fear', one single attitude towards the relevant situation happens to be retained. But this does not mean that the occupation of the subject's attention is not going through changes more or less constantly. Following Lambie & Marcel (2002), we may specify certain broad varieties of change, such as, e.g., change in what they call the 'focus' of attention. In fear-episodes, the focus the subject's attention may shift, for instance, from 'world-focus'–e.g., focus on the nature and the gravity of some threat, or focus on its perceptible fearsome characters–to 'self-focus'–e.g., focus on one's specific vulnerabilities. Furthermore, it may shift from focus to either the threat or the vulnerability to focus on action– i.e., one's coping potential–as, for example, when one looks for hideouts or possible escapes, either literal or metaphorical.[151] Do such shifts look like they are governed by association or inclination? When contemplating the nature and gravity of a threat leads us to having mental images and thoughts bearing on our vulnerability or our coping potential, we remain well within limits imposed by our being interested in the situation *as* fearsome, or dangerous, or threatening. All such images and thoughts are relevant to the reigning attitude toward the situation. Although they are distinct 'objects' of attention, they are objects of the same global interest. Let me elaborate a bit on this point.

Evans (1970, §4.2) in his very interesting discussion of the variety of attention he calls 'interrogative', describes a situation arguably analogous in structure to the one at issue here. He characterises the state of consciousness the subject is in while engaged in this form of sustained attention as one in which a single 'master-idea' is organising the subject's consciousness around a complex, thematic 'object of attention':

[151] Lambie & Marcel (2002) provide an overview of many more modifications in 'focus' and 'mode' of attention. They assume that during emotion attention may be shifted onto any emotional component. Arguably, however, not all such modifications are compatible with continued emotional experience. A telling example is provided shortly hereafter. Cf., Frijda (2005) for discussion.

> Interrogative attention is typically the attention manifested in problem-solving. In a problem-solving context there may be a number of factors that are thought to be relevant to the solution of the problem. When these simultaneously occupy attention *there is no question of the investigator being distracted by them*. He will only feel distracted by things that have no relation to the object of enquiry. Let us define as a 'relevancy system' the sum of those considerations that are taken by the attender to have a bearing on his investigative task. A 'divided attention' can then be said to be one in which the objects of attention do not both belong to a reigning relevancy system. When that happens we have a conflict between objects of attention. (op. cit., 118-119, italics added).

Emotional situations are not unlike Evans' *problem-solving* situations: the emoting subjects are motivated to cope or deal with particular situational demands, and, to do so, the occupation of their attention must be constrained so that only mental images and thought relevant to the their current interest–viz., Evans' 'master-idea'–shall tend to occupy their attention.[152] Assessing the nature or gravity of a threat, one's vulnerability to it, or one's coping potential are all relevant to dealing with 'fearsome' situations, and, hence, there will be no question of the emoting subjects being *distracted* by shifts in focus like the ones that I have mentioned. Conversely, emotion is easily punctured by shifts of attention leading to objects that do *not* belong to the system of relevancy that they, so to say, 'set up for us.' A very telling example of how clearly demarcated the boundaries of relevance are in the case of emotion is distraction due to a shift in focus from emotional situations as evaluatively connotated to the way in which one's body feels. Although somatic feelings are emotional components, and although, arguably, they contribute importantly to the felt character of emotion, they are not possible objects of affective interests–unless the eliciting situation is precisely how one's body feels. Therefore, they cannot become objects of attention in sustained emotional engagement. When one is in the grip of fear, as Deonna & Teroni (2012) put it, "the body is

[152] Teroni (forthcoming) discusses the relevance of Evans' concept of a 'master-idea' to the way in which emotions structure consciousness.

felt as a gestalt of bodily sensations which consists in being ready to respond to the object in a certain way", not, by contrast, as an atomistic "constellation of variously located bodily feelings." (87). Frijda (2005) illustrates the disruptive effect of shifts of attention onto somatic feelings in the passage I quote below.

> In self-focus, analytic attention reduces felt bodily engagement to just that. Felt impulse to shrink back from a threat is transformed into felt muscle tension, just as the feeling of pointing can be transformed into feeling one's finger stretched. One comes to feel dizzy, one's heart racing, instead of feeling anxious or upset. […] Analytic isolation of information sources robs an experience of its emotional character." (op. cit., 482).

Let me now consider multi-attitude emotion episodes. If we consider how one's attitude may shift from being *ashamed* of one's bad drunken behaviour to *regretting* one's decision to drink so much and then *dreading* the prospect of facing the witnesses of the scene in the future, again, the transitions do not look like they are due to associations and inclinations. During the course of such an episode, intuitively, one becomes vividly aware of a 'precise' stretch of one's life spanning the (shameful) behaviour, the (regretful) decision, and the (dreadful) prospect in question. From the point of view of the subject, the episode is an ongoing process *rationally* unfolding over time as distinct aspects of the significance of the relevant situation come into view and gradually compose a more and more complex overall appraisal. The distinct objects of the specific 'episodic' emotions that one may want to *extract* from such episode are closely related to each other as elements of one single system of relevancy– i.e., they are all relevant to the ongoing exploration of the significance of one particular situation Each shift in attitude, at the moment at which it takes place, may perhaps be taken to mark the end of such process of exploration: the subject has now settled on one specific evaluation of the situation. The model I have in mind here is that of the relations between

'deciding' that something is the case–in this case, settling on an evaluation–and the ongoing process of 'trying to make up one's mind'–in this case, as regards the significance of a given situation (O'Shaughnessy's 2008, §17.3). However, despite the structural analogies, emoting subjects are not *trying* to settle on an evaluation, and the formation of their specific emotion attitudes does not mark the end of the process, for the latter is essentially open-ended. Rather, the emoting subjects' affective interest in emotional situations may evolve into specific variants that succeed one another in driving the ongoing process. In this case, as in the 'fear' case, the affective interest provides for thematic continuity, and the attitudinal transitions, like the shifts in focus, are internally rationally intelligible. To wit, it is very odd to suppose that the subject is *distracted* from her shame by her subsequent regret and dread. As Evans suggests, the level of complexity of single systems of relevancy is no threat to continuity of attending: "[o]n the contrary, the greater the extent […of the] system, the greater will be the concentration of attention. The threat to attention from *extraneous* elements […] will then be all but non-existent." (op. cit., 123). This leads to my final point, which regards the bearing of the fact that we have the capacity to engage in so-called emotion-*regulation* strategies on the idea that emotional self-awareness is purposive self-awareness. In striving to distract ourselves so at to terminate our emotion by interrupting an ongoing process of appraisal, as well as in striving to reappraise emotional situations so as to change our attitude, we manifest a keen awareness of which changes in the occupation of attention may lead to continuation *vs* interruption of engagement.[153] Now, all this is not to say that association does not play any role here. As Wollheim (1999) has emphasised, imagination and association are much present in emotion: "[a] common role that thought plays within emotion is to cement those identifications or mergings of one figure with another, and of one fact with another, which give the life of the emotions their chaotic, overgrown, vegetal

[153] On distraction see Verduyn *et al* (2015, 333) and Freund & Keil (2013). On reappraisal see Gross (ed.) 2007. McRae *et al.* 2010 for discussion of the difference between distraction and reappraisal as regulation strategies.

character." (117). However, despite its 'chaotic' and 'vegetal' character, emotional cognition exhibits a thematic continuity and diachronic intelligibility–across mono-attitude and multi-attitude episodes of emotion –that idle attending lacks. Relatedly, its 'fragility' in the face of shifts of attention leading *outside* the limits set by current affective interests contrasts with daydreams' distinctive 'flexibility'.

3 Explorative absorptions

Let me now turn to the respects in which emotional cognition is a *special* kind of sustained attending. In what follows, I will identify two non-sufficient, but very plausibly necessary characters of sustained emotional engagement. Much of what I will say articulates intuitions that will only be fully vindicated in the next chapter, by examining (positive and negative) affective interest and the way in which they relate us to value, but the intuitions can be made quite clear already. The two characters I have in mind, as anticipated, are 'open-endedness' and 'non-instrumentality.' I will consider them in turn.

By 'open-endedness', I want to capture two closely related aspects of the self-awareness associated to emotional cognition: (*i*) the emoting subjects' purpose is typically that of exploring the relevance to one's concerns of emotional situations, and not that of coping or dealing with them by attaining what Frijda's calls 'relational purposes'–e.g., in the 'fear' case, achieving one's own inaccessibility–and (*ii*) sustained emotional engagement does not have completion conditions. In chapter 1, I suggested that we may naturally understand Frijda's concept of 'readiness' as referring *primarily* to the phenomenon of interest. The rationale for the suggestion was that this idea suits an account of emotion episodes aiming to range over mono-attitude and multi-attitude emotion episodes. Interest, as Frijda suggests, is present in all emotion episodes, and arguably lies beneath the action tendencies differentiating the various 'episodic' emotions. I wish to suggest that understanding emotional action readiness

as referring primarily to interest is crucial if we are to capture the subject's point of view on their own engagement. From a mere third-personal or functional-biological perspective, emotional action readiness is viewed as 'motivating' and 'preparatory': emotion episodes impel and (mentally) prepare the emoting subjects to achieve certain relational purposes. Now, if we think of such motivation and preparation in terms of the arousal of an interest in a situation *as* concern-relevant, we can do justice to varieties of emotional cognition and behaviour that do not lend themselves to third-personal treatment. In other words, we may account for the fact that some, if not much, emotional cognition does not *seem* to be directed at achieving relational purposes–or not from the subject's point of view. Let me illustrate. Although it is biologically compelling that one of the functions of fear is to motivate and prepare us to achieve our inaccessibility, not all fear-related (mental) activity has internal intelligibility in light of such purpose. It is true that during fear-episodes we often think of and look for hideouts and escapes; however, we may also entertain thoughts about the threat and the nature and extent of our vulnerability to it that do not bear *means-to-end* relations to relational purposes. For example, our attention may be captured by the sharpness of the teeth of the quickly approaching Rottweiler, and we may imagine how easily they would sink into our legs and anticipate the pain we may experience. None of this seems instrumental to achieving relational purposes. An account of emotional cognition must accommodate such cognitive activities, for, arguably, they are the most 'emotional' ones: they characterise the experience of one who really is in the grip of fear as opposed to one who is coldly assessing a threat so as to take some measures. If all this sounds at least initially plausible in the case of fear, it will be even more convincing if we consider emotion attitudes whose motivating dimension is not so clearly geared toward a course of action, such as Tappolet's (2014, ch.2), *contemplative* emotions. Tappolet takes this sub-class of emotion attitudes to encompass, e.g., 'awe', 'relief' and 'admiration', and any episode of emotion elicited by fictions or artworks. But the idea may be extended to real-life emotions that happen to be elicited in situations in

which no course of action is available or in which the costs of acting outweigh the potential benefits–consider, for instance, cases in which we have non-epistemic reasons not to be angry with our boss or our partner. Frijda (2004) nicely puts this point: "[t]he link between emotion and action is intimate; *yet it is weak*. Anger has intimate links to aggression, but few angers […] go that far" (163, italics added).

Let me now turn to non-instrumentality. Let me begin by noting that during the course of episodes of emotion, we enjoy a certain degree of non-distractibility: we are typically *absorbed* in our engagement. Now, in general, if we manage to become absorbed in an ongoing activity, we are then somehow oblivious to the purpose which explains our original engagement; therefore, the activity cannot be taken to be sustained by an awareness of its instrumentality to attaining such purpose. A we will see shortly hereafter, the purpose remains the underlying motivation, but our absorption lends its own drive to the activity. In emotion, the situation is similar in this respect in that we are typically unaware of the concerns that underlie our emotional reactions–this is at least part of the so-called 'opacity' of emotions[154]–so that we cannot appeal to an awareness of the purpose to explore how our concerns might be at stake to explain our engagement. This is also true of cases in which we can easily learn which concern is at stake, for instance, in perceptual fear-episodes. In such episodes, we do not seem to deploy our attention as we do because we take that to be conducive to assessing whether and how our concern for our physical integrity is at stake. Arguably, if we found ourselves in a burning building, not being trained fire-fighters, our attentional engagement would be driven by an interest in the situation as 'scary', 'horrific' or 'fearsome' as opposed to an interest in the situation as *merely* 'dangerous', i.e., as being a threat to our physical integrity. A Deigh (1994) puts it, "something that *looks* dangerous is something one can *infer* is dangerous from the way it looks, whereas one need make no inference to see that

[154] For extended discussion of the opacity of emotions with respect to underlying concerns see Teroni (2016).

something looks scary." (842, italics added). When in the grip of emotion, we do not typically make such inferences. Therefore, sustained emotional engagement is not plausibly *sustained* by the product of such inferences, at least not typically and not in the most distinctive cases. The absorption which characterises emotional cognition is closely related to the way in which sustained emotional engagement is *interest* driven; there is a sense of 'interest' which conveys specifically affective absorption, as opposed to absorption which constitutes the success of 'trying to focus'–that is, to achieve a certain degree of non-distractibility. In the next section, I will conclude this chapter by discussing in more detail unemotional absorptions so as to prepare the ground for discussing its emotional counterpart. I will then come back to the suggestions I have made about the relation that the latter bears to value in the next chapter.

4 Absorption and intrinsic interests

The concept of 'interest' has three main usages: it may be used in an abstract manner, as a placeholder for whatever reason or motive lies behind one's attending to an item or topic,[155] it may be used to refer to the state one achieves when one's attempts to become absorbed are successful, and it may be used to refer to a *sui generis* motivational phenomenon: roughly, the state we are in when we are absorbed in an activity because it is interesting, i.e., because of some kind of 'appeal' the activity has. In the next chapter I will focus on this third sense of the term, and I will try to make room for negative affective interests by spelling out the concept of a 'negative' appeal. I will conclude this chapter by discussing the other two senses of the term so that we can put them aside. To illustrate the contrast between non-affective

[155] In this, it is similar to Davidson's (1980) 'desire' or 'pro-attitude' although its breadth is arguably narrower. 'Desire' or 'pro-attitude' covers all those things which might, jointly with beliefs, rationalise action or activity, including affective episodes and whims and fancies (Johnston 2001, 187-188), 'interest' instead covers stative and more stable pro-attitudes–e.g., all our standing concerns–and seems especially apt to convey our 'ultimate' concerns as well as what we may call 'remote' or 'ulterior' purposes–i.e., purposes achieved by means of given actions in virtue of somewhat incidental connections. For example, it would be odd to explain one's pouring oneself a glass of water every hour by appealing to an 'interest' in drinking, as opposed to wanting or desiring to drink every hour, but one may remove this oddity by replacing 'drinking' with, e.g., 'keeping hydrated', i.e., by mentioning an explanatory standing concern.

and affective interests consider the following. One may explain one's dedication to one's daily piano practice, for example, by mentioning one's standing interest in earning a diploma, or making one's parents proud, or becoming more 'well-rounded', and one may do so while admitting to one's lack of *interest* in music for piano. Analogously, one may explain why one is devoting so much time to reading all the classics of western literature by mentioning an interest in becoming well-read, while admitting to finding the classics boring and *uninteresting*. By contrast, one may explain one's daily piano practice and one's reading of the classics–with no need to invoke further reasons or motives–simply by saying that one finds these activities *intrinsically* interesting. So far, I have reserved '*interest*' in italics, to this *sui generis* motivational phenomenon. Let me turn to non-affective interests.

In the sense in which 'interest' is a placeholder for whatever reason or motive lies behind one's sustained attentional engagement with an item or topic, the suggestion that our sustained emotional engagement is *interest*-driven would be trivial, for whenever we keep our attention fast onto some item or topic for some time we must be doing so for some reason, that is, *in view of* an interest that we are currently pursuing. Even if one were to maintain one's attention onto something having no reason other than just preventing it from spontaneously wandering–as one may do, for example, in (successful) meditation–and even assuming that whatever it is that one maintains one's attention onto in such practices–for example, one's breath, an object, a mantra etc.–is not *interesting* for one, one may still appeal to the concept of 'interest' to do such explanatory work. For example, one may appeal to an interest in calming one's mind so as to counter anxiety, or just to an interest in verifying for how long one can maintain one's attention onto an item. In any case, the attending the subject is engaged in, while it is not, *ex hypothesis*, intrinsically interesting, still counts as driven by an interest: what motivates the subject's sustained attending is the likely causal instrumentality of the latter to the obtaining of the state of affairs the subject is interested in. Now, while pursuing an interest by means of sustained attending, in given circumstances, we may end up

absorbed in the activity in a way that makes the interest somewhat redundant motivationally. While one's engagement is still, in some sense, motivated by such extrinsic interests, it is now continuing "by its own drive", as Woodworth (1918, 72) puts it. Woodworth nicely illustrates this pervasive attentional phenomenon in the passage I quote below.

> Often voluntary effort is needed in order to get a task started, to overcome repugnance, inertia, and distracting influences. The extraneous motive brings the horse to the water, but real drinking does not occur except from thirst, that is to say, from a *desire for the particular results obtained by the activity in progress*. […] It is only when an activity is running by its own drive that it can run thus freely and effectively; for as long as it is being driven by some extrinsic motive, it is subject to the distraction of that motive. […] It is not true, then, that the motive that initiates a given activity furnishes motive force for the whole activity; it simply leads the performer up to the act, but the motive force for the act itself must be inherent." (70-71, 1918, italics added).[156]

So, for example, if the initial motivation to start practicing meditation is to make one's anxiety subside, this interest will plausibly motivate the reiterate *efforts* that are required to keep one's attention fixed on the object, that which typically involves patiently bringing one's attention back to the object every time one notices that it has wandered. 'Successful' meditation, however, involves becoming absorbed in the object (or the activity) in a way that renders effortful exercises of the will to stay on topic more and more sporadic, and eventually unnecessary. The same goes, it seems, for any case of trying to focus or to become absorbed in an activity. Now, in such cases we wouldn't always want to say that the subject has acquired an *interest* in the activity in the sense that they are now motivated by some kind of appeal the activity now has. Rather, we seem to have this very basic capacity to become absorbed in what we do. This sort of absorption is *just* the success of trying to focus.[157] James seems to

[156] Cf. Allport's (1937) concept of 'functional autonomy.'
[157] Cf. O'Shaughnessy (2000, 386), who argues that this is the *only* sense in which one may be *trying* to attend, for it is the only sense in which one may properly *fail* to do so.

confirm this idea in an argument aimed to show that *all* instances of sustained attending involve this sort of absorption. He also implicitly suggests an explanation of this very basic capacity as one that exploits the natural tendency of the attention to wander to *new* things. In the passage I quote below, James is introducing a distinction between what he calls 'immediate' attending and 'mediated' or 'voluntary' attending. The claim he wants to make is that "we never make an effort to attend to an object except for the sake of some remote interest which the effort will serve" (416) and that sustained attention essentially involves the acquisition of what he calls a 'passive' interest allowing the activity to continue with no need for effortful exercises of the will.

> *There is no such thing as voluntary attention sustained for more than a few seconds at a time.* What is called sustained voluntary attention is a repetition of successive efforts which bring back the topic to the mind. The topic once brought back, if a congenial one, *develops*; and if its development is interesting it engages the attention passively for a time. […]. This passive interest may be short or long. As soon as it flags, the attention is diverted by some irrelevant thing, and then a voluntary effort may bring it back to the topic again; and so on, under favorable conditions, for hours together. During all this time, however, […] it is not an identical *object* in the psychological sense, but a succession of mutually related objects forming an identical *topic* only, upon which the attention is fixed. *No one can possibly attend continuously to an object that does not change*. (422-423, italics original).

James' 'passive' interest captures the second sense of the term which would make the claim that sustained emotional engagement is *interest*-driven rather trivial. We may notice in fact that explanations in terms of James' 'passive' interest and explanations in terms of 'remote' or extrinsic interests are not exclusive, for one may engage in an activity in pursuit of a remote interest and then develop James' 'passive' interest as the activity begins to continue by its own drive and exercises of the will to bring the topic back to mind are no longer necessary. Still, the original, remote interest remains the underlying source of the subject's motivation

as shown by the fact that when the passive interest 'flags', and attention wanders, an exercise of the will is required to re-engage. In the next chapter, I will differentiate the absorptions distinctive of sustained emotional engagement from absorption as the success of an attempt to become absorbed. This will clarify the above suggestions about the non-instrumentality of sustained emotional engagement in connection to how we relate to value.

Chapter 6
EMOTION, INTEREST AND VALUE

Introduction

I have argued that sustained attentional engagement driven by intrinsic interest is still too broad a notion to capture sustained emotional engagement. One may manage to become absorbed in activities such as counting the lines in some pattern, reading medical leaflets, solving logical puzzles, rehearsing, or executing dancing routines, while not being emotionally motivated to do as one does. All that is required is that one becomes oblivious to the extrinsic interests that originally motivate such activities. An empirical hypothesis we may come across in the field of so-called 'positive' motivational psychology is that successful absorption in ongoing activities is inherently rewarding, that is, it always involves some kind of affect.[158] This hypothesis fits well, and seems further elucidated by, the hypothesis that amongst the various standing concerns which make up our evaluative sensitivity is a standing concern for our good functioning (Frijda 2007, 138-141). The rationale for positing the concern, as I have already mentioned,[159] is that minimal interferences with ongoing activities tend to cause negative affect, even when the activities are apparently *insignificant*.[160] Conversely, it is not

[158] Cf. Deci & Ryan (1985) on so-called intrinsic motivation.
[159] See §3.7.
[160] Cf. Frijda (2007): "Recall the example of frustration when one's pencil point breaks during writing. It angers. Evidently, smooth continuation of an ongoing act represents a concern. The issue is not that one's writing will be finished half a minute later, but that pursuing one's active intent was halted before its time. Examples of a similar kind are when reaching for the salt cellar, say, and someone snatches it away when you're on the point of grasping. Even when he snatches it to hand it over to you, you feel bad. Your hand floats idly in space,

implausible to suppose that we undergo positive affective changes whenever we function especially well and the activity advances 'smoothly.' This might nicely explain what exactly absorption, or the acquisition of an intrinsic interest contributes *motivationally* such that only if we are intrinsically interested reiterated exercises of the will become dispensable. However, even if this were the case for all instances of sustained attending, including, for example, reading medical leaflets and counting the lines in some pattern, that is, even if it could be shown that affect *always* plays some role in sustained attention, we should arguably see this as some kind of pleasure in activity, not to be confused with other kinds of affect and certainly not emotion. To see this, first of all, consider that in such cases intrinsic interest and absorption always precede affect. Secondly, such cases are all cases of positive reward. Finally, and more importantly, the positive reward derives from absorption itself as the success of trying to focus–and not the *appeal* of the very activity in which we are engaged, therefore, arguably, such cases are not even identical with positive affective interest.

I will start to elucidate the concept of affective interest by arguing that active states of affective interest in a given activity obtain whenever we prefer being engaged in that activity rather than not and when such preference is grounded in (our acquaintance with) the appeal of the activity. The simplest and most intuitively compelling examples of this are provided by cases of *enjoyment* in (or of) our attentive activities as forms of positive affective interest. In enjoyment, unlike unemotional absorptions, we are not just absorbed, or oblivious to whatever extrinsic interests, if any, lie behind the activity, but our engagement in our activity is driven by the appealing character of the activity in ways which make extrinsic interests explanatorily and motivationally *inert*. I will then show how this may extend to negative affective interests by introducing the notion of *negative* appeal. Finally, I will turn to the distinctive immediate intelligibility that positive and negative interests seem to confer to

stupidly. Or when you are angrily voicing a reproach, but the other person apologizing before you have finished. Or when, in the dark, your foot moves for the next step on the stairs but you are already at the top." (ivi, 138).

emotionally motivated cognition and behaviour. To do so, I will elaborate on Johnston's (2001) analysis of affective motivation. Johnston views affect as a *sui generis* motivational phenomenon,[161] which makes affectively motivated cognition and behaviour *immediately* intelligible to the agent and "apt or fitting in a way that silences any demand for justification." (189). This view is developed by Johnston as part of an argument in defence of what he calls Detectivism, the view that values are out there to be detected as opposed to being projected by emoting subjects onto a neutral world. As we will see, Johnston's view readily accounts for positive emotional motivation. However, it needs to be further developed if it also is to account for negative emotional motivation. There are in fact structural disanalogies between being emotionally motivated to do x by x' *positive* appeal and being emotionally motivated to do x by x's *negative* appeal.

In order to draw out the specificity of this view of affective motivation, I will then contrast it with the one put forward by Brady (2013). Brady rejects Johnston's proposal and contends that affect far from silencing any demand for justification, motivates us precisely to seek evidence bearing on whether affectively motivated cognition and behaviour are warranted, i.e., reasons and evidence that might (dis)confirm our emotional appraisals. I will suggest that Brady misses Johnston's point about affect's power to silence any demand for justification. Anticipating a bit, Brady takes Johnston's view that affect makes our acting and desiring intelligible to us in a special manner to imply that emotional experiences are *sufficient* reasons for our evaluative judgements as cognitive elements mediating between affect and our intentional behaviour. Johnston's main point, however, is that affect makes our actions and desires *readily* intelligible to us, that is, with no need for intermediary evaluative judgements.

[161] Johnston uses the term 'affect' to pick out his target, namely "a direct experience of the appeal of things" (182), and so will I, but he makes it clear that he is exclusively concerned with *emotional* affect as opposed to affect involved in other affective reactions. As far as I can see, this interpretation is not undermined by his clarification that he will not deal with *the* emotions–viz., specific emotional attitudes–which he takes to "typically arise after one is drawn to or repelled by something" as a result of a process which crucially involves "consciousness of one's position vis à vis the repellent and the appealing" (footnote 1). All this entails is that he takes affect or affective engagement to occur prior to the acquisition of specific emotional attitudes.

I will thus reject Brady's positive account of affective motivation by arguing that the sort of immediate intelligibility Johnston identifies is a crucial plausibility requirement on any such account, and this regardless of our view of the epistemic *credentials* of emotional experience.

1 The structure of positive and negative affective interests

It is part of the idea that affective reactions motivate us to act and deploy our attention the way they do–i.e., in an impelled or coerced manner–that *when* we undergo affective reactions–i.e., at the time–we *just* prefer being engaged in the affectively motivated mental activity rather than not: when engaged in emotional cognition, instrumental–i.e., extrinsically motivated–exercises of the will are not required to keep us on our topic. However, it is crucial for this view to be able to get off the ground that we show that we may, in this sense, prefer being engaged in given activities, not only when we *enjoy* such activities, but also when they feel to be called forth by negatively appraised emotional situations. In other words, it is crucial that we make room for the notion of negative alongside positive affective interests.[162] Let me start with positive affective interest.[163] The existence of close links between enjoyment and absorption has been stressed by a number of authors, among which, Ryle (1954):

> Now to say that someone has been enjoying a smell or a walk at least suggests and maybe even implies that he has been interested in the smell or in the exercise and the incidents of the walk–*not* that he gave his mind to them in e.g., the *sedulous* way*,* but rather that *his mind was taken up by them in a spontaneous way*. (Ryle 1954, 142–3, italics added)

[162] This proposal is Brentanian in spirit in that it elaborates on the idea that all affective, and more generally, motivational phenomena–i.e., phenomena that are neither presentations nor judgements of presentations–are phenomena of (positive or negative) interest. Cf: "The term 'interest' is usually used only to designate certain acts which belong to the class we are describing, namely, those in which curiosity or inquisitiveness are aroused. Yet *it is not inappropriate to describe every pleasure or displeasure in something as interest*, and every wish, every voluntary decision is an act of taking an interest in something, too." (PES, 153 [II, 35], quoted in Textor 2017, 229, italics added). The same intuition seems to be at play in Frijda's suggestion that emotions are 'variants of interest'.

[163] In what follows, I rely heavily on Textor's (2017) discussion of the nature and structure of enjoyment. Thanks to Fabrice Teroni for pointing out to me the pertinence of this literature and for helpful discussion.

In the above passage, by being 'interested', Ryle clearly means being intrinsically interested or absorbed: the contrast between being interested in something and '*giving* one's mind' clearly matches James' previously examined distinction between 'immediate' and 'voluntary' attending. Ryle goes on to suggest that enjoyment *is* this kind of attending. As Textor (2017) observes, however, "Ryle's proposal is [as regards enjoyment] too wide; he has not managed to give a feature that is distinctive of enjoyment, but of a broader notion, namely that of taking an interest." (228, italics added). This is shown by the fact that there are clear cases in which an agent takes an interest in an activity without thereby enjoying it.

> [B]eing absorbed in an activity cannot be the same as enjoying it. Parking a car in a tight space might be so difficult that I need to devote all my attention to it. For instance, I am reluctant to stop parking the car because if I take my mind off it, I will not succeed. But I don't enjoy parking the car! Similarly, alarming things may capture one's attention. If an angry Rottweiler approaches me, I will be fully immersed in watching it and nothing may distract me. Yet I don't enjoy watching it. (op. cit., 229).

Now, Textor's examples seem sufficient to deny that enjoyment and absorbed attending are coextensive, let alone identical, but they 'overshoot': we do not need to look at cases in which we are 'positively' *not* enjoying the activities to reject Ryle's proposal. As we have seen, it is quite possible for us to become absorbed in an activity neither appealing nor unappealing–such as, e.g., reading medical leaflets or meditating. As I have mentioned, it is possible that engaging in such instances of sustained attending in the required manner–i.e., without needing to remind oneself of one's *underlying* extrinsic interests–is sometimes facilitated by some kind of affect, namely, the reward one derives from one's very absorption as an instance of good or smooth functioning. This kind of affect, however, is arguably not the same as that involved in enjoyment or positive affective engagement: the activity in question

is not in itself appealing to us and our engagement, accordingly, is not felt to be impelled or coerced to any degree. The examples that Textor offers, I suggest, are better suited to illustrate the contrast between positive and negative emotional absorption. Both being absorbed in parking the car in a tight space and being absorbed in watching a Rottweiler approaching, are quite naturally described as emotional situations. For example, I may be *anxious* that I may not be able to park the car right now and I may be *afraid* of the angry Rottweiler. If I am anxious about my being unable to park the car, say, because I am late and cars are in line behind me waiting for me to park, I will be emotionally motivated to park the car, and this will lead to absorption in the task. This case may help clarify the sense in which emotional attending is often taken to be 'effortless.' Parking a car in some tight spot may be difficult and may thus require some *effort*. To insist on the effortless character of emotional attending is not to insist on the fact that our attentional engagement may not involve efforts; rather, it is to insist on the fact that one is to an extent impelled or coerced to devote one's attention to the task at hand, so that no effortful exercise of the will is required to keep one's mind onto the task. One had achieved a certain degree of non-distractibility. In this negative emotion appears to be rather similar to enjoyment. To further illustrate the similarity, we may contrast these situations with analogous albeit unemotional ones, or to make the contrast starker, with situations in which, while I have deliberately set for myself the task of parking the car or watching the dog, I am currently affected by other things. Let's suppose, for instance, that I am watching the dog's behaviour because my friend has asked me to do so, and that, while I have chosen to indulge him and promised to do so, I am not myself afraid of the dog, since it does not seem erratic to me. Or else imagine that I am parking the car knowing well that the space is very tight and that the operation requires my full attention, but I am emotionally drawn to the voice from the car radio because my match is on. In both cases, although I do, in the end, give my mind to the tasks and succeed to become absorbed, my absorption is still motivated instrumentally: I am not affectively interested in the tasks,

but the latter are instrumental to the satisfaction of extrinsic interests or concerns–such as, e.g., reassuring my friend and not damaging my car. In other words, I have reasons to try to become absorbed by the tasks I have set for myself. To the extent that I do succeed, I qualify as intrinsically interested, but this does not mean that my intrinsic interest is affective: my activity is not felt to be urgent or coerced; it is still fully up to me, and although I am oblivious to my extrinsic interests, they are not made explanatorily and motivationally inert *from my point of view*; they still remain the source of my motivation.

At this point, one may raise the following objection to the claim that positive and negative affective interests lead to the same kind of absorption. While in positive affective interest we do not feel the need to further inquire into whether the subject has some extrinsic interest–"I'm enjoying this" is a perfectly satisfying answer to the question "Why do you do this?"–in negative affective interest, it seems that we do feel this need, and it also seems that the agent may be plausibly expected to be able to mention some extrinsic interest. For example, being asked why she is so absorbed in parking the car, the subject might reply that she is so because she doesn't want to have to look for another parking spot, or because she doesn't want the people in the cars in line behind to wait for her. Analogously, being asked why she gives her mind to the angry Rottweiler, the subject might reply that she wants to be able to anticipate, and hence protect herself from, the dog's attack. In what sense, then, is negative affective interest an intrinsic interest? As we have seen in the previous chapter, intrinsic interest and extrinsic interest are not mutually exclusive concepts. This seems to also apply to cases in which the intrinsic interest is affective. For example, one may *enjoy* one's solfège exercises while, at the same time, being extrinsically interested in exercising because of its likely instrumentality to gaining perfect pitch or making one's parent proud. My suggestion is that while in the case of unemotional absorptions the extrinsic interests remain the source of the motivation, in enjoyment and emotions, the mere availability of extrinsic interests does not make the agent's engagement any less affectively motivated: the engagement is *not* driven

by extrinsic interests which provide motivational support for deliberate exercises of the will to give one's mind to the task. There is, however, a structural *disanalogy* between positive and negative affective interest lying behind the above objection. In enjoyment, I prefer an activity *A* because I positively appraise *A*. In negative affective interest, I prefer an activity *A* because I negatively appraise a situation *S* that *A feels* to cope or deal with. This does not mean that I positively appraise *A*–e.g., I do not enjoy parking the car and I do not enjoy watching the Rottweiler, just like I do not enjoy entertaining 'painful' mental images and thoughts while in the grip of regret, or shame, or jealousy. In other words, in both cases I may be said to prefer *A*, but in positive affective interest I prefer *A* because I like or love *A*, whereas in negative affective interest I prefer *A* because I dislike the emotional situations *A* is felt to deal or cope with. In the next section, I will elaborate on this and try to draw out the specificity of negative emotional motivation by discussing the relations between negative affective interest and value.

2 Johnston on affective motivation and evaluation

In this section, I will introduce Johnston's arguments for the view that emotional motivation differs from purely *rational*–viz., extrinsic–motivation in that it makes affectively motivated cognitions and behaviour intelligible to the subject in a distinctively 'immediate' way, that is, with no need for mediating reason-based evaluative judgement. Johnston's (2001) purpose is to defend the view that value is there in the world to be detected as opposed to being projected onto neutral 'objects' and show that affect or affective engagement is necessary in order to access or 'disclose' such values.[164] It is important to understand what Johnston means by 'value.' He clarifies that he is concerned with "inherently sensuous" values, i.e., "the *utterly determinate* versions of such determinables as the beautiful, the charming, the erotic

[164] On the notion of 'disclosure of value', see Vannello (2018).

(in the narrower sense), the banal, the sublime, the horrific and the plain old appealing and the repellent." (182). These values are called inherently sensuous to advert to the fact that they would remain inaccessible to subjects deprived of the appropriate "sensibilities": "[t]hought and judgement directed at these […] values could not be generated simply by the understanding," Johnston suggests, for "[s]omething akin to sensing and sense-based imagination is required to make them available as topics for thought and judgement." (ibid). Restricting the scope of the claim that affect allows us to access values to inherently sensuous values entails two things. Firstly, the values at issue cannot be so-called 'response-*independent*' evaluative properties,[165] such as the 'dangerous' for fear or the 'incongruous' for amusement, since even though a *special* kind of understanding of the 'dangerous' and the 'incongruous' would arguably remain inaccessible absent appropriate sensibilities (Deonna & Teroni 2012, 123-124), acquiring the concept of such a property doesn't require being 'sensitive' to it (Deigh 1994, 840-843): we only need to be able to assess threats and relative vulnerabilities and identify incongruities. The values at issue must be 'response-dependent', for example, the 'fearsome' and the 'amusing', and more generally, the 'appealing' and the 'repellent'. Secondly, Johnston is concerned with our access to the utterly *specific* way in which such evaluative properties are instantiated by particulars or situations. Acquaintance with the way in which emotional situations exemplify such evaluative properties, according to Johnston, essentially requires affect or affective engagement:

> Seeing the utterly specific ways in which a situation, animal or person is *appealing* or *repellent* requires an appropriate affective engagement with the situation, animal or person. Absence of appropriate affect makes us *aspect-blind*. The world then appears more neutral than it is, and our immediate evaluational thought and judgement becomes impoverished. *Intrinsic motivation is lost, and eventually our own ongoing*

[165] Cf. Dokic & Lamaire (2013); Prinz (2004, 60-66); Tappolet (2014, chs.2-3).

> *activity lacks a kind of ready intelligibility, which the giving of reasons to ourselves hardly makes up for.* (op. cit., 181, italics added).

In view of my more limited purpose, I can afford to bracket the question whether evaluative properties are essentially detectible or projected and focus on the relation between what Johnston calls intrinsic motivation and the ready, immediate intelligibility which characterises affectively motivated (mental) activity. Getting a clear grasp on this connection will put us on the right track towards identifying what I take to be an essential plausibility requirement on an account of sustained emotional engagement that may capture the emoting subject's point of view on their own mental activity. Johnston's claim that without appropriate affects our intrinsic motivation is lost is connected to the claim that "the presence of the affect can make the desire or action especially intelligible to the agent himself. It can make the desire or act seem apt or fitting in a way that silences any demand for justification." (ibid., 189). Jonhston illustrates this claim by means of an analogy with perceptual experience

> Perceptual experience makes certain immediate perceptual beliefs about the perceived scene seem apt or fitting in a way that silences any demand for justification for those beliefs. If the same beliefs were to immediately arise in the absence of perceptual experience then they would then lack a certain intelligibility from the inside. […] In the same way affective disclosure of sensuous goods makes desire *readily intelligible from the inside*. Were a desire to come upon one without one's being in some way taken with the desired object or outcome then there would be a lack of intelligibility in one's having the desire. The question of justification–"Why am I doing or wanting this?" would naturally arise. (op. cit., 189-190, italics added).

He supports this proposal by considering a case in which the subject's ongoing activity is unintelligible to the subject himself and arguing that this lack of intelligibility is due to the

absence of appropriate affective engagement as opposed to absence of an appropriate reason-based evaluative judgement.

> I could find myself with a very strong urge to turn the knobs of doors counterclockwise, even though I have no interest in opening doors and no interest in feeling the solidity and texture of the knobs in my hand. The urge could be very effective, but it would not render intelligible what I am doing, least of all to myself. I aim to argue that the urge does not rationalise or render intelligible my action because, as stipulated, it involves no affect, and so no disclosure of the appeal of what it is an urge to do. (op. cit., 190).

So, the knob-turner has an urge to turn knobs, and he acts in the way that he does motivated by this urge. *Ex hypothesis*, the agent has no interest whatsoever in the activity: "no interest in opening doors and no interest in feeling the solidity and texture of the knobs in my hands." Johnston maintains that in such a case, although "the urge could be very effective, […] it would not render intelligible what I'm doing, least of all to myself". How can we diagnose this sort of lack of intelligibility? Johnston proceeds by opposing two main options: absence of relevant affective engagement, and, hence, absence of felt appeal of the activity–namely, the option he favours–and second, absence of the relevant evaluative judgement, that is, deeming the activity somehow worthwhile. The first case he considers is one in which an agent positively evaluates a course of action but cannot get himself to engage in it because affect is lacking. His example is that of a man who wants to put on weight by eating protein bars, but, although, *ex hypothesis*, he takes the action to be worthwhile, the protein bar is so unappealing to him that he cannot get himself to do so. This, in contrast to the knob-turner's condition, is a perfectly intelligible condition to find oneself in "even from the inside" (ibid.). Of course, Johnston continues, there is a competing explanation of the man's failure to act, for the case of the man who wants to put on weight but cannot get himself to do so by

means of eating the bars is also analysable as weakness of the will understood as failure to be motivated to act by one's best evaluative judgement. Analogously, the knob-turner's behaviour may be taken to lack some kind of intelligibility from the inside since the agent has no tendency to positively evaluate the way he is impelled to act. Johnston thus attempts to distinguish the internal intelligibility he is after from the sort of intelligibility that may be contributed by reason-based evaluative judgement. He invites us to imagine that the man gets himself to eat the bar by sheer willpower and he suggests that the act might still lack some kind of intelligibility from the agent's perspective, for even though the act is positively evaluated, "[t]he thought "Why on earth am I doing this?" can coexist with judging that there is *something* to be said for eating the Power Bar. For I can still fail to see anything appealing about the Power Bar" (op. cit., 191). I find that this sort of example is intuitively compelling. There is a kind of intelligibility from the inside which characterises cognitions and behaviour motivated by affect that evaluative judgement cannot replace. The immediacy of this sort of intelligibility is connected to the concept of intrinsic interest: if my motivation to act is due to the instrumentality of the act to some outcome I positively evaluate, then my act is only intelligible to me in virtue of my awareness of such instrumentality. By contrast, if I act in the way I do because the act is *appealing*, i.e., with no extrinsic interests in view, then, the act is readily intelligible.

The issue I want to address is the issue of whether we can elicit the same intuition for negative affective interests. Although Johnston explicitly declares to be concerned with determinates of the 'appealing' and the 'repellent', his discussion is in fact restricted to the 'appealing.' Disclosure of the 'repellent', however, because of the structural disanalogy between positive and negative affective interests I have highlighted, does not just repel. To wit, not all forms of emotional motivation involve avoidance and very few involve cognitive avoidance. Rather, Johnston's 'repellent', and more generally, negative affective appraisal confers a negatively

appealing character to the affectively motivated cognition. Over the next section, I will try to substantiate this suggestion.

3 The negatively appealing

Johnston considers a possible explanation of the knob-turner's behaviour that may introduce some kind of intelligibility from the inside. The knob-turner's urge may be construed on the model of 'needs' or 'wants' understood as motivation to act so as to alleviate the discomfort of deprivation. The idea is that perhaps the knob-turner's urge is like a hunger or *craving* in that the agent feels discomfort if she does not engage in it and is therefore impelled to get out of her state by doing what her state is an urge for, i.e., knob-turning. Now, in so far as the cessation of previous discomfort is essentially relieving, we may imagine the agent *enjoying* satisfying her urge to some extent. Johnston maintains that the action would still lack the intelligibility he is interested in: being motivated to satisfy an urge in the sense above clarified, that is, in virtue of a dynamic between the suffering of want or *deficit* and the pleasure of relief coming from satisfaction of an urge can coexist with still failing to see the appeal of the act itself.[166] The idea here is that one does not *enjoy* the activity, just like one does not enjoy parking the car under the impulse of anxiety or monitoring the behaviour of the Rottweiler under the impulse of fear. What is still missing is, according to Johnston, a *positive* affect or affective engagement with the act and, hence, disclosure of its appealing character.

While this seems correct, it is not clear that lack of positive affective engagement makes the desire and the action unintelligible *in the same sense* in which munching on the bar was found to be unintelligible. The thought "why on earth am I doing this?" does not seem to arise in this case. I am acting this way because, although the act is not positively appealing, I'd still rather act as I do than not. The act is negatively appealing to me. Let me elaborate on this.

[166] For an analysis of this sort of interplay see Ombrato & Phillips (2020). The following discussion owes much to exchanges with Edgar Phillips.

To clarify, what makes the act appealing is not the pleasure that one derives from relief. Johnston does recognise that deprivation may enhance appetite, so that, for example, food that I do not consider especially appealing if fully satiated may become extremely appealing if I am very hungry. To this one may add that it is not clear that we can fully detach our motivation to get out of an unpleasant state of want from our motivation to get into the more pleasant state of relief. In other words, one may construe Johnston's example as involving disclosure of positive appeals. To do this, however, would miss Johnston's point: there is an important difference between being affectively engaged with something in a way that reveals that something's appealing character and being motivated to get out of an unpleasant state and into a more pleasant one, and Johnston suggests that only the former involves affective disclosure. The rational for this restriction is that in seeking the pleasure of relief we are focussed *inwardly*, in a way that Johnston takes to be incompatible with our being really taken with the source of our relief as opposed to enjoying how it *makes* us feel. While I agree with this, Johnston's way of rejecting the 'urge-like' explanation seems to reveal a view of *urges* that, I suggest, may blur the differences between sense-based and drive-based affective reactions, on the one hand, and emotional reactions, on the other. On this view all negative affective reactions can motivate us merely by making us wish to rid ourselves of an unpleasant sensation. Let me begin by noticing that the sort of motivation associated to cravings like 'hunger' is sometimes to be construed as involving a disclosure of the repellent character of one's deprived situation, and not merely inward focus on the negative feelings associated to such state of deprivation. Arguably, in the case of physical craving, this sort of motivation does not strictly belong to the craving; rather, it belongs to emotional reactions the craving may trigger, for example, distress or anguish about one's felt need. Johnston recognises this emotional dimension–or better emotional enrichment–in the case of positive sensory affects by granting that sensory pleasure and pleasure in activity may disclose positive value *if* we are engaged with the object or activity and not focused inwardly on the way they

make us feel (cf. op. cit. 200-203). Now, if we recognise the structural disanalogy between positive and negative affective interest as well as the similarity between the motivational situations to which they give rise, we may accommodate the intuitive differences between the two and still ascribe to them the same kind of internal intelligibility. The 'urge-like' explanation of the knob-turner's behaviour, I am suggesting, can be construed as involving disclosure of the *negative* appeal of turning doorknobs. Thus construed, it may in principle involve the same kind of intelligibility Johnston ascribes to positive interest. Johnston rejects the 'urge-like' explanation by assimilating the behaviour with pathological compulsive behaviour, such as the urge of pathologically compulsive agents to wash their hands under the assumption that argue that from the agents' point of view their acts might seem absurd and lack a certain intelligibility–e.g., the agents might know full well that her hands are clean and there is no real need to wash them again and again. However, is this kind of lack of intelligibility the same Johnston highlights in the previous example? It seems plausible that if the action lacks *some* intelligibility, but this has to do with the knowledge I just mentioned: the agent knows that her action is not worthwhile, she doesn't positively evaluate it and she doesn't identify with it. But the unavailable intelligibility is in this case is the intelligibility afforded by positive evaluative judgement and not the one afforded by affective engagement. At the same time, the action might be *immediately* intelligible to her, she'd rather wash her hands than not because doing so feels to correct or remedy a situation she appraises negatively, that is, whose repellent character is (apparently) revealed or disclosed. Why such engagement, thus construed, would be any less affective than positive affective engagement is not clear. Johnston does consider the proposal that the lack of intelligibility might derive from lack of *identification* with the act, but dismisses it by considering cases in which an addict who doesn't identify with her addictive behaviour still finds it intelligible:

> I might be an unwilling addict, ruining my life with endless chess games played over the Internet. I repudiate my desire to play chess all the time, I want to get on with my life, but my chess playing is at least intelligible to me. I am really taken with chess, its appeal is manifest to me. My desire to play chess all the time is not alien to me in the way my urge to turn knobs is. My problem is that I cannot make effective my higher-order desire to forget chess and get on with my life. This may be partly because my first-order desire is full of a vivid sense of the appeal of chess. As a result, my acting on my first-order desire to play chess is intelligible to me, in contrast to my knob-turning. Higher-order desire, or whatever the apparatus of identification comes to, does not make for this difference. (192)

It is not clear why addictive chess-playing should be any less intelligible in the sense Johnston is after than compulsive knob-turning is. The relevant difference in the way the examples are devised is that in the chess example, the agent is "really taken with chess"–he *likes* or *loves* chess and "its appeal is manifest to [him]". But part of what it means to be addicted to doing something is that refraining from that something is negatively affectively appraised, and if we take this to involve a disclosure of that something's *negative* appeal, the two cases seem to be essentially analogous. The intelligibility Johnston is after seems to inhere in the ready availability of *motivation* afforded by affective interests: when I'm really taken with an ongoing activity there is no need to remind myself of the reason why I'm doing what I'm doing, for the reason is right there; it is my immersion in the very activity I'm engaged in generated by the appeal the activity has for me

> One is so taken with what one is doing that there is no attention left over to notice or think about anything irrelevant, including irrelevancies like verbalized positive evaluations of what one is doing. Consciousness of oneself as set in a world of opportunities and obstacles to one's will disappears, and the sense of time becomes distended. The ongoing activity is experienced as so appealing that we are willing to do it for its own sake, without regard for its consequences, and even when it is difficult or dangerous. (ibid., 193)

My suggestion is that the same ready availability of *motivation* is afforded by negative interests. Analogously, my absorption in watching the angry Rottweiler, while not positively appealing, as opposed to, say, a bird-watcher's absorption, is compelled in a way that silences the demand for any justification. Watching the dog, while not enjoyable, is (negatively) appealing; the appeal is due to the fact that my engagement feels to deal somehow with situations whose 'scary' or 'fearsome' character is disclosed to me. This is very different from its being due to the fact that it feels to remedy or alleviate the unpleasant sensations associated with emotional suffering or discomfort. I am not focused inwardly, and I'm really taken it with my ongoing activity. In the next section, I will conclude by examining Brady's (2013) view of affective motivation and I will contrast the general picture of the emotional subjective perspective emerging from such view with the one I have been developing.

4 Emotions, evaluation and the emotional subjective point of view

According to Brady, emotional experience, far from silencing any demands for justification, motivates us precisely to seek reasons and evidence bearing on whether our initial, reflexive emotional reactions are actually justified. In what follows, I will reconstruct his argument and I will argue that it rests on two interpretive confusions. I will then turn to the prospects of Brady's positive account of emotional motivation. Let me start by considering Brady's reconstruction of Johnston's analogy between emotional and perceptual experience.

> Although Johnston talks about affect making desire or action intelligible to the agent, I assume that, given the analogy with perceptual experience, he also thinks that affect makes evaluative belief or judgement intelligible to the agent. Indeed, this would have to follow if we think that desire and action are often consequent upon evaluative belief, so that the belief is the intermediary element between affect and intentional behaviour. If so, then Johnston can be understood as holding that emotional and

> perceptual experiences are sufficient reasons for evaluative belief or judgement. A reason counts as sufficient in so far as it is good enough to justify some action or belief by itself; a sufficient reason makes action or belief permissible for the subject, we might say. (Brady 2013, 84).

Now, although the view that affect makes immediately intelligible certain evaluative beliefs is attributable to Johnston–one may issue evaluative judgments such as, e.g., that something is banal or sublime, based on affective disclosure of such values, and such judgement would according to Johnston be made readily intelligible by affect–the view that affective desiring and affectively motivated behaviour require evaluative beliefs as *intermediary* elements is not. Johnston's proposal is precisely that desire and action based on affect are immediately intelligible and that this intelligibility, absent *current* affective engagement, cannot be derived from mere evaluative judgements, however 'genuine' (Johnston 2001, 193). Brady seems to misconstrue the analogy in that he takes it to hold with respect to the strength of the epistemic *credentials* of the two. According to him, the analogy 'overstates' the epistemic credentials of affect and emotion. To oppose this, Brady argues that it is simply not true that we take our emotional reactions to give us sufficient reasons for evaluative judgements in the way in which we take perceptual experience to give us sufficient reasons for perceptual descriptive judgements. To see this, we just need to observe how different our epistemic situation as perceivers is from our epistemic situation as emoting subjects: on the one hand, in normal circumstances–i.e., absent epistemic defeaters–it would be strange if "after citing the fact that I have a visual experience as of a red car outside my flat, I was then asked for, or felt the need to discover, further reasons or evidence for this belief" (Brady 2013, 86), while, on the other hand, "we often *feel the need* to discover reasons or evidence in support of our emotional experience in normal circumstances, and hence in the absence of defeaters." (ibid.). Brady offers the following example to illustrate this felt need.

> Consider one's experience of fear, upon hearing a noise downstairs as one is trying to get to sleep. Let us stipulate that circumstances are normal: our hearing is good, we have not been taking hallucinogenic drugs, etc. In this situation, we are typically motivated to seek out and discover additional reasons or evidence. In particular, we are motivated to seek out and discover considerations that have a bearing on whether our initial emotional "take" on the situation, namely that we are in danger, is accurate. We strain our ears to hear other anomalous noises, or rack our brains trying to think of possible non-threatening causes for the noise. We do not, in other words, think that the demand for justification is silenced, or take the representational content of our emotional experience as true by default. It would be a strange and paranoid person indeed who took each and every such feeling of fear as a sufficient reason to believe that he was in danger. (op., cit., 86-87)

We may distinguish two main issues with Brady's interpretation of Johnston's argument: firstly, the relevant analogy holds with respect to the kind of intelligibility that emotional and perceptual experience may contribute and not with respect to epistemic strength of the reasons they might provide. Johnston's proposal is that emotional experience provides a sort of intelligibility "the giving of *reasons* to ourselves hardly makes up for." (op. cit., 181). Johnston's point is thus that if certain desires where to arise *absent* affective engagement with their objects, and, hence, absent affective disclosure of value, these desires would be devoid of a certain kind of ready intelligibility from the inside. For example, if one were to desire someone, in the narrowly erotic meaning of the term, absent affective engagement, i.e., absent affective disclosure of his or her erotic appeal, then, the desire would be unintelligible. In this respect, the analogy with perception seems to hold, for if one came to have certain beliefs about sensible qualities of an object without perceptual experience of that object, then, such beliefs would be devoid of intelligibility from the inside. Johnston makes the point more salient by imagining someone who suddenly acquired "the gift of imageless clairvoyance" and the epistemic situation of such person *vis à vis* their descriptive judgments

based on their clairvoyance (op. cit., 189).[167] Secondly, and relatedly, if in the above quote we replace 'dangerousness' with 'scariness' or 'fearsomeness', the intuition that Brady tries to elicit about the felt need to confirm our emotional appraisals becomes much more elusive: the idea that during fear-episodes we do not take the way we feel at face value in that we are motivated to assess whether the situation is scary or fearsome does not have initial traction. In general, the less one can resort to alternative evidence for the instantiation of a given property, the stronger the epistemic credentials of one's source of evidence are, and while evidence that something is dangerous is obtainable in a number of ways, fear seems to be the sole source of access to the utterly determinate manner in which my hearing a noise as I am about to fall asleep is scary or fearsome. Thus, even if the relevant respect of the analogy with perceptual experience were strength of epistemic credentials, Johnston's argument would not be so easily undermined.

Although Brady's objections to Johnston seem to miss their mark, Brady's own positive account of emotional motivation may still be taken to capture a genuine element of the phenomenon, for active cognitive engagement with emotional situations of the sort Brady highlights does occur sometimes during the course of emotional reactions. According to Brady, emotional experience is typically accompanied by cognitive processes that are aimed at (dis)confirming our initial emotional reactions: emoting subjects do not tend to take the way in which things are evaluatively presented to them at face value; rather, they are prompted to establish whether they have reason or evidence that may back up their emotional appraisals. My suggestion is that even though assessments of the credentials of our emotional appraisals do sometimes accompany emotional experience, for we may doubt whether things stand as our emotional experience seems to reveal, our engagement in such assessments cannot be taken to be what the emotional experience itself motivates us to do.

[167] A real-life example is arguably provided by the epistemic situation of *blindsight* patients vis à vis their correct guesses on what is presented in their blind perceptual hemi-field. See Smithies (2011).

Let me start by noting that the approach to emotion I have been defending is similar to Brady's in one respect: he too adopts a diachronic point of view on emotional reactions and highlights the importance of their persistence for our understanding of their nature and function. Moreover, he too connects emotional persistence to attentional persistence and the latter to *active* explorations of the value of the emotional situations.

> Emotions such as fear and anger stay with us; they are not simply short-term reflexive interruptions to our mental life, but often persist and dominate that life so that we remain focused on and attentive to danger, infidelity, and slights. Think again, to illustrate, about what it is like when one is awoken, in the dead of night, by a noise downstairs and as a result becomes afraid. In normal circumstances one's fear is not over very quickly; rather, one remains in a fearful state as one listens attentively for further noises, tries to think of possible non-threatening explanations, rehearses strategies for dealing with the potential danger, considers possible escape routes, and so on. These are the cognitive *accompaniments* to something else that persists, namely the mobilization of one's bodily resources which constitutes one's being primed or prepared to fight or flee. […]. Similar points can be made about anger, resentment, sadness, shame, guilt, love, and many other emotions. (op. cit., 92).

On Brady's account, however, the relations between emotional and attentional persistence are not as intimate than the ones I have articulated. Brady considers attentional persistence a typical accompaniment of emotional reactions, not an element of their causal structure. Accordingly, the active exploration of the value of emotional situations on which Brady focuses is not part of the emotional reactions as complex cognitive-affective episodes; it is not an element of sustained *emotional* engagement. We may think of the sort of cognition Brady is interested in as deliberate emotion-regulation, that is, the sort of reappraisal that we seek when we take a step back from our initial emotional reactions so as to favour a change in attitude. Brady thus takes our exploration of emotional situations to be caused by an extrinsic interest in assessing whether our initial emotional reactions are justified as opposed

to being driven by affective interests. Now, while it sounds plausible that emotional persistence occasions and facilitates Brady's explorations of emotional situations, that this is what emotions themselves motivate us to do, let alone essentially motivate us to do, is not. As I have argued, the emotional influence on attention is analogous to that of other kinds of affective reaction: we feel as if we won't be able to engage in any other activity requiring attentional resources until we have answered the demands of the relevant eliciting situations, that is, until we have dealt with such situations as evaluatively connotated in the way our emotional appraisals 'take' them to be. The claim that emotions essentially motivate the search for evidence bearing on their own justification is no more plausible than the view that pain essentially motivates us to uncover evidence bearing on whether some body part requires tending and that thirst and hunger essentially motivate us to discover whether we are dehydrated or undernourished. Of course, there are many disanalogies between pain, thirst, hunger etc. and our emotional reactions: while the former are our primary sources of information about tissue damage, dehydration and undernourishment, emotional reactions are not the primary sources of our apprehensions of objective relational properties such as 'danger', 'sleight', 'infidelity', 'noxiousness', 'degradation' etc. Moreover, while sense-based and drive-based affective reaction are *typically* extremely reliable sources of such information, emotional reactions are not. Our evaluative sensitivities are clearly susceptible of being better attuned or refined over time, but they are also susceptible of being corrupted in many ways. Accordingly, unemotional active assessments of our emotional appraisals *may* accompany emotional experience, that is, we may *doubt* that our emotional appraisals match how things evaluatively stand, and we may thus seek further evidence. This does not mean, however, that our engagement in such assessments is motivated by our emotional reactions themselves. It is motivated by our doubt. Furthermore, the scope of Brady's account of our active cognitive engagement with emotional situations is arguably rather limited. Although in the case of fear we have an appealing candidate for the response-independent evaluative

property the actual instantiation of which emoting subjects may seek to assess, since the link between fear and dangers, or threats is often transparent to emoting subjects, this is not the case for most emotion-attitudes (cf. Dokic and Lamaire 2013, 239-241). Awareness of this sort of link is sometimes well-established in adult and articulate individuals for many of the emotions that they may feel. An awareness of this sort may in such individuals play the role that Brady wishes to assign to the way in which emotions relate us to value, i.e., prompting and guiding a search for reasons and evidence. But our awareness of such links remains a matter of discovery, and if we soon discover the link in the case of fear, this is not the case for the link between disgust and the 'intoxicant', shame and the 'degrading' or amusement and the 'incongruous'.[168] If we are to account for emotional cognition in its variety and aiming to capture what is most distinctive in it, we need to think of it as driven by affective interests in emotional situations, as characterised by what I have called their exploratory, open-ended character and their *non-instrumental*, absorbed character. Now, one key theoretical advantage of the approach to emotion that I have articulated in this thesis as regards our understanding of the emotional subjective perspective is that it allows us to build *active* cognitive engagement with emotional situations *into* the emotion episodes *vs* treating it as an accompaniment. Therefore, it may capture the agential aspects of the emoting subjects' *emotional* relation to values, i.e., the dependency of their apprehensions of value upon an ongoing active process of exploratory absorption. The emoting subjects' commitment to given, thick evaluations–i.e., their emotion attitudes–is not a commitment the *fixity* of which is in need to be remedied by means of unemotional engagement like that which takes centre stage in Brady's picture. In virtue of its dependency on an ongoing active appraisal, such evaluative commitments are dynamic and 'tentative', as is clearly shown by the great diachronic complexity of our most articulate multi-attitude emotion episodes.

[168] Cf. Dokic & Lamaire (2013): "[t]he wide disagreement […] about the proper pattern of appraisal specific to each emotion is in and of itself evidence that there are no response-independent evaluative properties that we all 'see' as belonging to the objects of our emotions" (240).

The main motivation behind the recent resurgence of interest in the concept of 'process', and more specifically, in deploying this concept in the analysis of various aspects of the mind, is the centrality of this concept to our understanding of the subjective point of view. As Stout (2018, 2) puts it, "the subjective perspective, if it has a special role in understanding a subject's mental life, is a perspective on their ongoing mental life–i.e., their life as a process." Now, on the one hand, the concept of 'process' is often taken to be crucial to our understanding of *agential* aspects of the mind and, one the other hand, emotions are one of the *paradigms* of passivity in the mental domain. However, as I have argued, emotion episodes are *both* episodes of affective reaction and episodes of our sustained emotional engagement: their mode of production crucially involves not only continual elicitation but also a *line* of agential-rational causation continuously perpetuated over time, a line of agential-rational causation to which they owe their thematic continuity and internal diachronic intelligibility. If this is correct, then, if we want our accounts of emotions to do justice to the emotional subjective point of view, we need to treat emotions as complex dynamic episodes over time. With this work I hope to have removed some obstacles to doing so and to have taken some steps in the desired direction.

Bibliography

Allport, G. W., (1937). 'The Functional Autonomy of Motives'. The American Journal of Psychology, 50(1/4), 141-156.

Bain, A., (1865). *The emotions and the will*. London: Longmans, Green, and Company.

Blackman, R. D. (2013). 'Intentionality and Compound Accounts of the Emotions'. *The Southern Journal of Philosophy*, 51(1), 67-90.

Brady, M. S. (2013). *Emotional insight: The epistemic role of emotional experience*. Oxford: Oxford University Press.

Brentano, F. (1995). *Descriptive Psychology*. London: Routledge.

Broad, C. D. (2014). *Five Types of Ethical Theory*. New York: Routledge.

Cannon, W. B., (1932). *The Wisdom of the Body*. New York: Norton.

Claparède, E. (1928). 'Feeling and emotions'. In M. L. Reymert (Ed.), *Feelings and emotions*. Worcester: Clark University Press.

Cooper, R. P. (2010). 'Cognitive control: componential or emergent?' *Topics in Cognitive Science* 2(4)), 598-613.

Colombetti, G. (2005). 'Appraising valence'. *Journal of Consciousness Studies*, 12(8-9), 103-126.

Crane, T., (998). 'Intentionality as the mark of the mental. In A. O'Hear, (ed.), *Contemporary Issues in the Philosophy of Mind*. Cambridge: Cambridge university Press.

Crowther, T. (2009). 'Watching, sight, and the temporal shape of perceptual activity'. *Philosophical Review*, 118(1), 1-27.

Crowther, T. (2010). 'X—The Agential Profile of Perceptual Experience'. In *Proceedings of the Aristotelian Society* (Vol. 110, No. 2pt2, 219-242). Oxford, UK: Blackwell Publishing Ltd.

D'arms, J., & Jacobson, D. (2000). 'Sentiment and value'. *Ethics*, 110(4), 722-748.

———————————(2003). VIII. 'The Significance of Recalcitrant Emotion (or, anti-quasijudgmentalism)'. *Royal Institute of Philosophy Supplements*, 52, 127-145.

———————————(2006). 'Anthropocentric Constraints on Human Value'. In R. Schafer-Landau (ed.), *Oxford Studies in Metaethics. Vol 1*. Oxford: Oxford University Press.

Dainton, B. (2000). *Stream of Consciousness: Unity and Continuity in Conscious Experience*. New York: Routledge.

Deigh, J. (1994). 'Cognitivism in the Theory of Emotions'. *Ethics*, 104(4), 824-854.

————(2001). 'Emotions: The Legacy of James and Freud'. *The International Journal of Psychoanalysis*, 82(6), 1247-1256.

DeLancey, C. (2006). 'Basic moods'. *Philosophical Psychology*, 19(4), 527-538.

Deonna, J., & Teroni, F. (2012). *The emotions: A Philosophical Introduction*. New York: Routledge.

De Sousa, R. (1987). *The Rationality of Emotion*. Cambridge, MA: MIT Press.

————————-(1990). "Emotions, education and time". Metaphilosophy, 21(4), 434-446.

Dixon, T. (2012). 'Emotion: History of a Keyword in Crisis'. *Emotion Review* 4(4),338-344.

Dokic, J., & Lemaire, S. (2013). 'Are emotions perceptions of value?', *Canadian Journal of Philosophy*, 43(2), 227-247.

Eilan, N., (1998). 'Perceptual intentionality. Attention and Consciousness'. In A. O'Hear (ed.), *Contemporary Issues in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Ekman, P., (1973). *Darwin and Facial Expression*. New York: Academic Press.

————————(1994). 'Moods, Emotions and Traits'. In P. Ekman & R. J. Davidson (eds.), *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press.

Ekman, P. & Davidson, R. J. eds., (1994). *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press.

Ellsworth, P.C. (1991). 'Some Implications of Cognitive Appraisal Theories of Emotion'. In K.T. Strongman (ed.), *International Review of Studies on Emotion*.

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*. Oxford: Oxford University Press.

Evans, C. O. (1970). *The Subject of Consciousness*. New York: Routledge.

Evans, G. (1985). *Collected Papers*. Oxford: Clarendon Press.

Faucher, L., & Tappolet, C. (2002). 'Fear and the Focus of Attention', *Consciousness & Emotion*, 3(2), 105-144.

Fredrickson, B. L., & Kahneman, D. (1993). 'Duration neglect in retrospective evaluations of affective episodes', *Journal of Personality and Social Psychology*, 65(1), 45.

Freund, A. M., & Keil, A. (2013). 'Out of mind, out of heart: Attention affects duration of emotional experience'. *Cognition & Emotion*, 27(3), 549-557.

Frijda, N., Mesquita, B., Sonnemans, J., & Van Goozen, S. (1991). 'The Duration of Affective Phenomena or Emotions, Sentiments and Passions'. In K. Strongman (ed.), *International review of Emotion and Motivation*.

Frijda, N. (1986). *The Emotions*. Cambridge: Cambridge University Press.

————(1993). 'The Place of Appraisal in Emotion'. *Cognition & Emotion*, 7(3-4), 357-387.

————(1994). 'Varieties of Affect: Emotion Episodes, Moods and Sentiments'. In P. Ekman & R. J. Davidson (eds.), *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press

————(2005). 'Emotion Experience'. *Cognition & Emotion*, 19(4), 473-497.

————(2007). *The Laws of Emotion*. Mahuah, NJ: Lawrence Erlbaum Associates, inc.

Gallie, R. (1954). 'Pleasure'. *Proceedings of the Aristotelian Society, Supplementary Volumes* 28, 147–63.

Gendron, M. (2010). 'Defining Emotion: A Brief History.' *Emotion Review* 2(4), 371-372.

Goldie, P. (2000). *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.

————-(2003). 'Emotion, Feeling, and Knowledge of the World'. In R. C. Solomon (ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions*. Oxford: Oxford University Press.

————(2009). 'Getting Feelings into Emotional Experience in the Right Way'. *Emotion Review*, 1(3), 232-239.

————(2012). *The Mess Inside: Narrative, Emotion, and the Mind*. Oxford: Oxford University Press.

Gordon, R. M. (1986). 'The Passivity of Emotions.' *The Philosophical Review*, 95(3), 371-392.

Grahek, N. (2001). *Feeling Pain and Being in Pain*. Oldenburg, Denmark: BIS-Verlag, University of Oldengurg.

Gray, J. R., Todd S. B., and Marcus E. R. (2002). 'Integration of Emotion and Cognition in the Lateral Prefrontal Cortex. *Proceedings of the National Academy of Sciences*, 99(6), 4115-4120.

Gray, J. R. (2004). 'Integration of Emotion and Cognitive Control', *Current Directions in Psychological Science* 13(2), 46-48.

Greenspan, P. S. (2014). *Emotions and Reasons: An Inquiry into Emotional Justification*. New York Routledge.

Griffiths, P. E. (1997). *What Emotions Really are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.

Gross, J. J., ed., (2007). *Handbook of Emotion Regulation*. London: Guilford Press.

Helm, B. W. (2001). 'Emotions and Practical Reason: Rethinking Evaluation and Motivation'. Noûs, 35(2), 190-213

Hornsby, J. (2012). 'Actions and Activity'. *Philosophical Issues*, 22(1), 233-245.

Intriligator, J. & Cavanagh P. (2001). 'The Spatial Resolution of Visual Attention'. *Cognitive Psychology* 43(3), 171–216.

Irving, Z. C. (2016). 'Mind-wandering is unguided attention: accounting for the "purposeful" wanderer'. *Philosophical Studies*, 173(2), 547-571.

Irving, Z. C., & Glasser, A. (2020). 'Mind-wandering: A philosophical guide'. *Philosophy Compass*, 15(1).

Irving, Z. C., & Thompson, E. (2018). 'The Philosophy of Mind-wandering.' In Kieran C. R. Fox & Kalina Christoff (eds.), *The Oxford Handbook of Spontaneous Thought*, 87-96. Oxford: Oxford University Press.

Izard C. E. (1971). *The Face of Emotion*. New York: Appleton (Century/Crofts).

————-(2010). 'More Meanings and more Questions for the term "Emotion"'. *Emotion Review*, 2(4), 383-385.

James, W. (1890/1950). *The Principles of Psychology*. New York: Dover

Jennings, C. D. (2012). 'The Subject of Attention', *Synthese* 189(3), 535-554.

Johnston, M. (2001). 'The Authority of Affect', *Philosophy and Phenomenological Research*, 63(1), 181-214.

Kagan, J. (2010). 'Some Plain Words on Emotion', *Emotion Review*, 2(3), 221-224.

Kenny, A. (1963). *Action, Emotion and Will*. London: Routledge.

Lambie, J. A., & Marcel, A. J. (2002). 'Consciousness and the Varieties of Emotion Experience: A Theoretical Framework', *Psychological review*, 109(2), 219-259.

Lazarus, R. S. (1991). 'Cognition and Motivation in Emotion', *American Psychologist*, 46(4), 352–367.

————-(1994). 'The Stable and the Unstable in Emotion', in P. Ekman, & R. J. Davidson, (eds.). *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press

LeDoux, J. (1996). 'Emotional Networks and Motor Control: A Fearful View', *Progress in Brain Research*, 107, 437-446.

Leeper, R. W. (1949). 'A Motivational Theory of Emotion to Replace "Emotion as Disorganised Response"', *Psychological Review*, 55, 5-21.

Lewis, M. D. (1996). 'Self-organising Cognitive Appraisals', *Cognition & Emotion*, 10(1), 1-26.

Marcel, A. J. (1993). 'Slippage in the unity of consciousness'. *Experimental and Theoretical Studies of Consciousnes*s, 174, 168-186.

Marks, J. (1982). 'A Theory of Emotion', *Philosophical Studies*, 42(2), 227-242.

Martin, M. G. F. (1997). 'The Shallows of the Mind', *Proceedings of the Aristotelian Society, Supplementary Volumes*, 71, 75-98.

———————-(2001). 'Out of the Past: Episodic Recall as Retained Acquaintance', In C. Hoerl, C. & T. McCormack (eds.), *Time and Memory: Issues in Philosophy and Psychology*. Oxford: Oxford University Press.

McDowell, J. (1996). *Mind and World*. Cambridge, MA: Harvard University Press.

McGinn, C., (2004). *Mindsight: Image, Dream, Meaning*. Cambridge, MA: Harvard University Press.

Meinong, A. (1917/1971). *On Emotional Presentation*. Translated by M. L. Schubert-Kalsi. Evanston: Northwestern University Press.

Milona, M., & Naar, H. (2020). 'Sentimental Perceptualism and the Challenge from Cognitive Bases', *Philosophical Studies*, 177(10), 3071-3096.

Mitchell, J. (2021). Affective Shifts: Mood, Emotion and Well-being', *Synthese*, https://doi.org/10.1007/s11229-021-03312-3.

———————-(2021b). *Emotion as Feeling Towards Value: A Theory of Emotional Experience*. Oxford: Oxford University Press.

Mole, C. (2011). *Attention is Cognitive Unison: An Essay in Philosophical Psychology*. Oxford: Oxford University Press.

Mole, Christopher, 'Attention', *The Stanford Encyclopedia of Philosophy* (Winter 2021), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2021/entries/attention.

Monsell, S., & Driver, J., eds. (2000). 'Control of Cognitive Processes', *Attention and performance XVIII*. Cambridge, MA: MIT Press.

Montague, M. (2009). 'The Logic, Intentionality and Phenomenology of Emotion', *Philosophical Studies*, 145(2), 171-192.

———————-(2014). 'Evaluative Phenomenology', In S. Roeser and C. Todd (eds.), *Emotion and Value*. Oxford: Oxford University Press.

Moors, A. (2009). 'Theories of Emotion Causation: A Review', *Cognition & Emotion*, 23 (4), 625-662.

Morris, W. N. (1989). *Mood: The Frame of Mind*. New York: Springer-Verlag.

Mulligan, K. (2004). 'Brentano on the Mind', In D. Jacquette, (ed.), *The Cambridge Companion to Brentano*. Cambridge: Cambridge University Press.

————————(2007). 'Intentionality, Knowledge and Formal Objects', *Disputatio* 2(26), 205-228.

————————(2009). 'On being Struck by Value: Exclamations, Motivations and Vocations', in B. Merker, (ed.), *Leben mit Gefühlen, Emotionen, Werte und ihre Kritik*. ISBN: 978-3-96975-106-0.

————————(1998). 'The Spectre of Inverted Emotion and the Space of Emotion', *Acta Analytica*, 89-105. Original French version: (1995) 'Le spectre de l'affect inverti et l'espace des émotions', in P. Paperman & R. Ogien, (eds.), *La Couleur des pensées*, *Raisons pratiques*, 6, 65-83.

Massin, O. (2020). 'Suffering Pains'. In D. Bain, M. Brady & J. Corns (eds.), *Philosophy of Suffering: Metaphysics, Value, and Normativity*, 75-100. New York: Routledge.

Melzack, R. (1975). 'The McGill Pain Questionnaire: major properties and scoring methods'. *Pain*, 1(3), 277-299.

Melzack R, Casey KL (1968). 'Sensory, motivational, and central control determinants of pain: a new conceptual model'. In Kenshalo D (ed) The Skin Senses. C. C. Thomas, Springfield

Mulligan, K., & Scherer, K. R. (2012). 'Toward a Working Definition of Emotion', *Emotion Review*, 4(4), 345-357.

Müller, J. M. (2017). 'How (not) to Think of Emotions as Evaluative Attitudes'. *Dialectica*, 71(2), 281–308.

————————(2021). 'The Spontaneity of Emotion', *European Journal of Philosophy*, 1-19.

Musil R., (1995). *The Man Without Qualities*. Croydon: Picador.

Naar, H., & Teroni, F., eds. (2018). *The Ontology of Emotions*. Cambridge: Cambridge University Press.

Na'aman, O. (2019). 'The Rationality of Emotional Change: Toward a Process View', *Noûs*, 55(2), 245-269.

————————-(2021). 'Emotions and Process Rationality', *Australasian Journal of Philosophy*, 99(3), 531-546.

Nash, R. A. (1989). 'Cognitive Theories of Emotion'. *Noûs*, 23(4), 481-504.

Neisser, U. (1976). *Cognition and Reality*. New York: W. H. Freeman & Company

Oatley, K., & Duncan, E. (1992). 'Incidents of Emotion in Daily Life', in K. T. Strongman (ed.), *International review of studies of emotion*, 2, 249–293.

Ombrato, M. D., & Phillips, E. (2020). 'The Mind of the Hungry Agent: Hunger, Affect and Appetite'. *Topoi*. 40, 517-526.

O'Shaughnessy, B. (2003). *Consciousness and the World*. Oxford: Oxford University Press.

Koralus, P. (2014). 'The Erotetic Theory of Attention: Questions, Focus and Distraction' *Mind & Language*, 29(1), 26-50.

Pashler, H., (1998). *The Psychology of Attention*. Cambdridge, MA: MIT Press

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). 'Attention and Performance'. *Annual Review of Psychology*, 52(1), 629-651.

Parkinson, B., & Manstead, A. S. R. (1992). 'Appraisal as a Cause of Emotion', in M. S. Clark (ed.), *Emotion*, 122–149.

Pessoa, L. (2009). 'How do Emotion and Motivation Direct Executive Control?', *Trends in Cognitive Sciences*, 13(4), 160-166.

————(2013). *The Cognitive-Emotional brain: From Interactions to Integration*. Cambridge, MA: MIT Press.

Peacocke, C. (1998). 'Conscious Attitudes, Attention and Self-knowledge', In B. C. Smith, C. Macdonald & C. Wright (eds.), *Knowing our Pwn Minds*. Oxford: Oxford University Press.

Pears, D. F. (1962). 'Causes and Objects of some Feelings and Psychological Reactions', *Ratio*, 4(2).

Peters, R. S., & Mace, C. A. (1961). Emotions and the Category of Passivity. Proceedings of the Aristotelian Society 62, 117-142.

Ploner, M., Freund, H.-J., & Schnitzler, A. (1999). 'Pain Affect without Pain Sensation in a Patient with a Postcentral Lesion', *Pain*, 81, 211–214.

Posner, M. I. (1980). 'Orienting of Attention', *Quarterly Journal of Experimental Psychology*, 32(1), 3-25.

Prinz, J. (2004). *Gut reactions: A Perceptual Theory of Emotion*. New York: Oxford university Press.

Reisenzein, R., & Schönpflug, W. (1992). 'Stumpf's Cognitive-Evaluative Theory of Emotion', *American Psychologist*, 47(1), 34.

Ryle, G. (1949/2009). *The Concept of Mind*. New York: Routledge.

Ryle, G. (1954). 'Pleasure', *Proceedings of the Aristotelian Society* 28, 135–64.

Roberts, R. (2003). *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press,

Roessler, J. (1999). 'Perception, Introspection and Attention', *European Journal of Philosophy*, 7(1), 47-64.

Robinson, J. (2005). *Deeper than Reason: Emotion and its role in Literature, Music, and Art*. New York: Oxford University Press.

————(2018). 'Emotion and Process', in H. Naar, & F. Teroni, (eds), *The Ontology of Emotion*, Cambridge: Cambridge University Press.

Ruckmick, C. A. (1936). *The Psychology of Feeling and Emotion*. New York: McGraw-Hill.

Russell, J. A. (2012). 'Introduction to Special Section: On Defining Emotion' *Emotion Review*, 4(4), 337.

Russell, J. A., & Barrett, L. F. (1999). 'Core Affect, Prototypical Emotional Episodes, and Other Things called Emotion: Dissecting the Elephant', *Journal of Personality and Social Psychology*, 76(5), 805.

Searle, J. R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT press.

Shand, A. (1914). *The Foundations of Character*. London: Macmillan.

Scanlon, T. M. (1990). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Scarantino, A. (2005). *Explicating Emotions*. Ph.D. diss., University of Pittsburgh.

————-(2012). 'How to Define Emotions Scientifically', *Emotion Review*, 4(4), 358-68.

————-(2014). 'The Motivational Theory of Emotions', in J. D'Arms, & D. Jacobson, (eds.), *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*. Oxford: Oxford University Press.

Scherer, K. R. (1987). 'Toward a Dynamic Theory of Emotion', *Geneva studies in Emotion*, 1, 1-96.

————-(2009). 'The Dynamic Architecture of Emotion: Evidence for the Component Process Model', *Cognition and Emotion*, 23(7), 1307-1351.

————-(2014). 'Affect bursts', in Stephanie H.M. van Goozen, Nanne E. Van de Poll, Joseph A. Sergeant, S.H.M. van Goozen, Joe A. Sergeant (eds.), *Emotion: Essays on Emotion Theory*. New York: Psychology Press.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford: Oxford University Press.

Scherer, K. R., & Wallbott, H. G. (1994). 'Evidence for universality and cultural variation of differential emotion response patterning'. *Journal of Personality and Social Psychology*, 66(2), 310.

Scherer, K. R., Wallbott, H. G., & Summerfield, A. B. (1986). *Experiencing Emotion: A Cross-cultural Study*. Éditions de la Maison des Sciences de l'Homme & Cambridge University Press.

Sherrington, C. (1900). 'Cutaneous Sensations', in E.A. Schäfer (ed.), *Textbook of Physiology*, London: Pentland.

Soteriou, M. (2013). *The Mind's Construction: The Ontology of Mind and Mental Action*. Oxford: Oxford University Press.

————-(2018). 'The Ontology of Emotion.' In H. Naar, & F. Teroni, (eds.) *The Ontology of Emotion*. Cambridge: Cambridge University Press

Sartre J. P. (1939/2002). *Sketch for a Theory of the Emotions*. Translated by Philip Mairet. Oxford: Routledge.

Smithies, D. (2011). 'Attention is rational-access consciousness.' In C. Mole, D. Smithies, W. Wu (eds.), *Attention: Philosophical and Psychological Essays*. Oxford: Oxford University Press.

Solomon, R. C. (1976). *The Passions: The Myth and Nature of Human Emotion*. Oxford: Anchor.

Steward, H. (1997). *The Ontology of Mind: Events, Processes, and States*. Oxford: Oxford University Press.

Steward, H. (2013). 'Processes, Continuants, and Individuals', *Mind*, 122(487), 781-812.

Stocker, M., & Hegeman, E. (1992). *Valuing Emotions*. Cambridge: Cambridge University Press.

Stout, R., ed., (2018). *Process, Action, and Experience*. Oxford: Oxford University Press.

Strawson, P.F. (1974). 'Freedom and Resentment', in P. F. Strawson, *Freedom and Resentment and Other Essays*. London: Methuen, 1–25.

Tappolet, C. (2014). 'The Normativity of Evaluative Concepts', in A. Reboul, (ed.), *Mind, Values, and Metaphysics*, 39-54.

————(2016). *Emotions, Values, and Agency*. Oxford: Oxford University Press.

_____(2018). 'The Metaphysics of Moods'. In H. Naar, & F. Teroni, (eds.) *The Ontology of Emotion*. Cambridge: Cambridge University Press

Teroni, F. (2007). 'Emotions and Formal Objects', *Dialectica*, 61(3), 395-415.

————(2016). 'Emotions, Me, Myself and I', *International Journal of Philosophical Studies* 24(4), 433-451.

————-(2017). 'Getting a Grip on Emotional Modes', in A. Cohen, & R. Stern, (eds.), *Thinking About the Emotions. A Philosophical History*. Oxford: Oxford University Press.

————(2018) 'Emotionally charged: The Puzzle of Affective Valence', in F. Teroni, C. Tappolet & A. Konzelmann Ziv (eds.), *Shadows of the Soul: Philosophical Perspectives on Negative Emotions*. New York, Routledge.

Teroni, F. (forthcoming). 'Affective Selves, Streams of Consciousness and Mental Time Travels', in I Velleman & V. Berninger, (eds.). *Philosophical Perspective on Memory and Imagination*. New York: Routledge

Textor, M. (2017). *Brentano's Mind*. Oxford: Oxford University Press.

Thalberg, I. (1964). 'Emotion and Thought', *American Philosophical Quarterly* 1,45-55.

Tomkins, S. S. (1995). *Exploring Affect: the Selected Writings of Silvan S Tomkins*. Edited by V. Demos. Paris: Editions de la Maison des Sciences de l'Homme & Cambridge University Press.

Tye, M. (2003). Consciousness and persons: Unity and identity. Cambridge, MA: MIT Press.

Vanello, D. (2018). 'Affect, Perceptual Experience, and Disclosure', *Philosophical Studies*, 175(9), 2125-2144.

Verduyn, P., Delvaux, E., Van Coillie, H., Tuerlinckx, F. and Van Mechelen, I. (2009). 'Predicting the Duration of Emotional Experience: Two Experience Sampling Studies', *Emotion*, 9, 83–91.

Verduyn, P., Van Mechelen, I., & Tuerlinckx, F. (2011). 'The Relation between Event Processing and the Duration of Emotional Experience', *Emotion*, 11(1), 20.

Verduyn, P., Van Mechelen I, & Frederix E. (2012). 'Determinants of the Shape of Emotion intensity profiles', *Cognition & Emotion* 26(8), 1486-1495.

Verduyn, P., Delaveaux P., Rotgé. J.-Y., Fossati, P., & Van Mechelen, I (2015). 'Determinants of Emotion Duration and Underlying Psychological and Neural Mechanisms', *Emotion Review* 7(4), 330-335.

Vuilleumier, P. (2005). 'How Brains beware: Neural Mechanisms of Emotional Attention', *Trends in Cognitive Sciences*, 9(12), 585-594.

Vuilleumier, P., Armony, J., & Dolan, R. (2003). 'Reciprocal Links between Emotion and Attention', *Human Brain Function* 2, 419-444.

Watson, D. (2000). *Mood and Temperament*. New York: Guilford.

Watzl, S. (2010). *The Significance of Attention*. Ph.D. diss. Columbia University. ProQuest Dissertation Publishing.

————(2011). 'Attention as Structuring of the Stream of Consciousness.' In C. Mole, D. Smithies & W. Wu (eds.), *Attention: Philosophical and Psychological Essays*. Oxford: Oxford University Press.

————(2017). *Structuring Mind: The Nature of Attention and how it Shapes Consciousness*. Oxford: Oxford University Press.

White, A. R. (1962). 'Attending and noticing', *Proceedings of the Aristotelian Society* 63, 103-126.

————(1964) *Attention*. Oxford: Basil Blackwell.

Williams, B. (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Wilson, J. R. S. (1972). *Emotion and Object*. Cambridge: Cambridge University Press.

Wollheim, R. (1999). *On the Emotions*. New Haven: Yale University Press.

Young, P. T. (1949). 'Emotion as Disorganized Response—a reply to Professor Leeper', *Psychological Review*, 56(4), 184–191.

Woodworth, R. S. (2018). *Dynamic Psychology*. New York: Columbia University Press.

————————-(1921). *Psychology: A study of mental life*. New York: Henry Holt and Co.

Wu, W. (2011). 'Attention as Selection for Action.' In C. Mole, D. Smithies, W. Wu (eds.), *Attention: Philosophical and Psychological essays*. Oxford: Oxford University Press.

————(2014). *Attention*. New York: Routledge.

Zajonc, R. 'On the Primacy of Affect', *American Psychologist* 39 (1984), 117–29.

Zamuner, E. (2015). 'Emotions as Psychological Reactions', *Mind & Language*, 30(1), 22-43.