

Joint aesthetic attention

Lucas Battich^{1,2}

¹Department of Philosophy, University of Tübingen, Germany

²Institut Jean-Nicod, EHESS, CNRS, École Normale Supérieure, PSL University, France

lucas.battich@uni-tuebingen.de

Forthcoming in *Synthese*
Author's version. Please cite published article.

Abstract

We experience and appreciate together with others the beauty of a sunset or an artwork, the sadness of a musical piece. Admiring the same aesthetic properties of objects together with others is a ubiquitous human activity. At the same time, it is commonly held that aesthetic properties constitutively depend on a subjective response. This brings up a particular challenge: How can we jointly focus towards the same aesthetic property when it is itself dependent on our separate responses? In this paper, I provide an analysis of joint aesthetic attention and explain how we can engage in it. I propose a multiple coordinative profile: co-attenders coordinate their perceptual attention to the object, but also coordinate the specific way they attend to it, and their specific responsive profiles. I advance a novel capacities-based probabilistic account of how this multiple coordination is possible.

Keywords: Aesthetic experience, Aesthetic properties, Joint attention, Collective attention, Aesthetic attention

1. Introduction

We enjoy together our favourite band's concert. We appreciate paintings together at the gallery. We are awestruck by a beautiful landscape while admiring it together with our loved ones. Enjoying the aesthetic properties of objects (expressive or formal properties; e.g., beauty, delicacy, power, balance, symmetry, and so on) together with others seems to be a ubiquitous human activity. But how does this happen? How do we jointly experience the same aesthetic properties of an artwork, a performance, or a landscape? In this paper, I introduce and analyse the notion of joint aesthetic attention. I advance an account of how it is possible to share our attention to the aesthetic properties of objects and scenes together.

The terminology of aesthetic attention can be ambiguous, depending on whether we emphasise the attentional *attitude* itself or the *target* of attention (e.g., aesthetic properties). This dual sense of understanding aesthetic attention points in turn to two different preliminary definitions of joint aesthetic attention:

Partial joint aesthetic attention (attitude): To engage in joint aesthetic attention is to jointly attend to the same object, event, or scene in a *specific aesthetic way*.

I call this *partial*, since we don't have to coordinate towards the same aesthetic property or have the same aesthetic experience, but we merely engage towards the same object in the same aesthetic way or with the same general aesthetic attitude. For instance, while in front of a Gentileschi at The Uffizi, we regard the painting aesthetically together, but I find it intense while you find it tepid, we are engaging in partial joint aesthetic attention. From the art gallery to the concert hall, partial joint aesthetic attention is a widespread phenomenon and, as it will become clear later on, a precondition to engage the full version:

Full joint aesthetic attention (properties): To engage in joint aesthetic attention is to jointly attend to the *same aesthetic properties* of an object, event, or scene.

During full joint aesthetic attention, we jointly focus our attention towards the same aesthetic property or have the same aesthetic experience. Here, we not only admire the painting together but, importantly, we jointly experience its intensity (or its crudeness, sadness, etc.). Here, it is often natural to use other terms than attention, and say, without losing sight of our phenomena, to jointly *perceive*, *appreciate*, or *experience* aesthetic properties, to see together the artwork's beauty or its crudeness, to hear together the sadness or the comedy in the music. Joint aesthetic experience might be a suitable alternative terminology, and I will also refer to it. In this paper, I focus mostly on full joint aesthetic attention (towards the same properties), since I believe it is the more challenging phenomenon, and will hence drop the "full" qualification unless needed. A bonus of the proposal is that it accounts for both versions in the same analysis.

This preliminary definition already brings to the fore two aspects to be developed. First, it refers to the ubiquitous phenomenon of joint *perceptual* attention, in which two individuals attend to the same perceptual object or event together, and they are aware of so attending (Bruner, 1974; Moore & Dunham, 1995). This mutual awareness of attending to the same object is what puts the "jointness" in joint attention. Suppose you are looking at the Gentileschi while I stand some distance away, looking at you and the painting. Here, we are both attending to the same painting, but you are not aware of that. This is a situation of merely parallel, but not joint, attention. I then move next to you, we exchange looks and gaze back to the art: we are now jointly attending to the painting. One challenge in analysing joint aesthetic attention consists of saying how it is possible to coordinate attention to the same aesthetic property so that we are mutually aware of doing so.

Second, the definition refers to aesthetic properties. Aesthetic properties are not, according to many accounts, on a par with garden-variety perceptual properties like shape or timbre. It is often held that aesthetic properties, along with other value qualities, are to some extent response-dependent properties (see, e.g., De Clercq, 2008; Johnston, 1989; Pettit, 1991). The *response-dependency thesis* holds that aesthetic properties depend constitutively or partly so, on the subject's attitude or response. For example, qualities such as beauty, elegance, and awfulness, require some sort of response from a human subject in order to be individuated or instantiated. An object being beautiful is, to some extent, for that object to give rise to certain emotional and/or evaluative responses in normal perceivers under normal circumstances (Schellekens, 2008).

Of course, many simple perceptual properties like colours and smells are also said to be response-dependent or secondary properties of objects. However, a key difference between secondary perceptual properties and aesthetic (and value) properties lies in the type of response. For perceptual properties like colour, response-dependence is typically said to be rigidly specified: the response is tied to the way in which physical stimuli cause certain sensory responses due to the physiology of our perceptual apparatus (López De Sa, 2013). In contrast, for value properties, including aesthetic ones, it is usually held that the required response is evaluative and/or affective, tied to human attitudes or judgments (Yates, 2008; Wright, 1992).

Here, I am concerned with more flexible forms of response-dependence, where the response is not rigidly constrained by direct perceptual sensation or by features of the sensory apparatus in typical healthy subjects¹. The relevant response may even be tied to more contingent conditions of the subject, like a particular cultural background needed to experience, say, a play as *tragic* in the classical sense. This type of flexible response-dependence may open the door to some kind of relativism² (López De Sa, 2013). The same object may be experienced as funny, tragic, elegant or graceful by *me*, but not so by *you*. It is precisely this possibility for relative responses that has led some philosophers to claim that we cannot enjoy aesthetic properties together at all. We are, Nick Zangwill argues, “in aesthetic solitary confinement” (2012, p. 384).

Any robust analysis of joint aesthetic attention should explain the phenomenon under a framework that endorses response-dependency about aesthetic properties. For, if the experience of aesthetic properties constitutively depends on the individual subjective responsiveness, how can we jointly attend to them and enjoy aesthetic experiences together? How do we coordinate our attention together towards an artwork’s beauty, the melancholy or happy character of a tune?

Previous discussion has focused on the possibility of joint aesthetic experience of music (Cochrane 2009; Polite 2019). However, the challenge of explaining aesthetic coordination under response-dependency remains. Besides encompassing a more general aesthetic domain than music, in this paper I propose an analysis of what joint aesthetic attention is, and explain how it is possible to engage in it. In what follows, I first advance a suitable framework of aesthetic experiences and a more specific definition of joint aesthetic attention (section 2). In order to explain how we engage in joint aesthetic attention, I introduce a novel capacities-based account, which lowers the conditions required for joint attention in general, including typical perceptual cases. Compared to other approaches, the explanatory focus here is on joint attention as a psychological capacity, involving a probabilistic coordination and recruiting possibly fallible estimates and assumptions (section 3). This results in a minimal but

¹ Some philosophers argue that some aesthetic properties are, like colour and shape, directly accessible by perceptual sensation alone (often termed *formal* properties), whereas many other non-perceptual aesthetic properties supervene on or are determined by these perceptual properties (Carroll, 2006; Hick, 2012). In section 5.1, I will address how my framework can account for Noël Carroll’s moderate formalism about perceptual experience and properties, which seems to reject the flexible response-dependency I assume here.

² Although it doesn’t directly imply anti-realism or subjectivism about aesthetic properties, as the same object may concurrently have several dispositional properties not presently being manifested through a subjective response (see section 5.2 below, and Schellekens, 2008).

explanatorily fruitful account, which can thus address the puzzle of joint aesthetic attention under the response-dependency thesis (section 4). After considering different aesthetic frameworks (section 5), I conclude with open questions about the possibility of collective aesthetic attention.

2. What is joint aesthetic attention

2.1. Aesthetic experience and response-dependency

Current theories of aesthetic experiences are a good starting point to get a better hold on joint aesthetic attention and how it negotiates the response-dependent issue. Joint attention to a target is, after all, a type of joint *experience* directed towards that target. Thus, I here assume that, at a suitable descriptive level, to jointly attend to the same aesthetic properties (e.g., sadness) of an object (e.g., a musical passage) is equivalent to jointly having the same aesthetic experience towards that object's properties (e.g. experiencing the musical passage's sadness). I here propose a simple account of aesthetic experience that can make good of the fact that it is directed towards a particular (response-dependent) aesthetic property in a particular object or event, and use this as a basis for an account of joint aesthetic attention.

Theories of aesthetic experiences can be roughly slotted into different families depending on where they locate the signature of the “aesthetic” (Nanay, 2018; Vernazzani, 2023). According to “structural” views, the “aesthetic” in aesthetic experiences is in the structure or quality of the experience itself rather than, for example, its content. Most structural views agree that aesthetic experiences require a particular attitude or response from the subject. Given this emphasis, they offer a suitable framework to build an account of joint aesthetic attention that can accommodate the response-dependency of aesthetic properties. From all structural formulations, I think that Jerrold Levinson's approach does the best job at neatly separating the different components of aesthetic experience, allowing to better pursue the issue of joint aesthetic attention. According to Levinson:

Aesthetic experience is experience involving aesthetic perception of some object, grounded in aesthetic attention to the object, and in which there is a positive hedonic, affective, or evaluative response to the perception itself and/or the content of that perception. (Levinson, 2016, p. 39)

By way of clarification, Levinson says that “aesthetic perception can be understood as the upshot of aesthetic attending” (2016, p. 39). It seems to me more parsimonious to replace “aesthetic perception” with perception simpliciter, while keeping the idea that such perceptual experience is grounded in or requires a specific aesthetic sort of attention. Building on (or rather scaling down) Levinson's view, I propose the following wireframe definition:

Aesthetic experience is experience involving:

1. perception of some object, event or scene;
2. grounded in aesthetic attention towards that object, event or scene; and
3. a particular responsive profile towards the perception itself and/or the content of that perception.

This wireframe definition is still flexible enough to accommodate different ways of filling up the details: what precisely aesthetic attention is, and what the character of the responsive profile is. For Levinson, aesthetic attention is “attention focused on an object’s perceivable forms and properties, for their own sake and in their full individuality, apart from the utility of so attending” (Levinson, 2016, p. 39). While Levinson doesn’t elaborate, Bence Nanay (2014, 2016) has recently proposed a more functional and tractable view of aesthetic attention, as being simply a special kind of perceptual attention, one that is both focused (vs distributed) with regards to objects and distributed (vs focused) with regards to properties: “our attention is focused on one perceptual object, but it is distributed among a large number of this object’s properties” (Nanay, 2016, p.13). More broadly, and given the popular Kantian notion that aesthetic experience involves a disinterested or detached attitude, with the object being appreciated or valued for its own sake, one could hold that aesthetic attention is just such an attitude.

The mere exercise of aesthetic attention so understood does not determine *which* aesthetic property is being experienced. The responsive profile condition is here the key to accommodate the response-dependency of aesthetic properties. As the subject’s response is constitutive of the character of aesthetic properties, it is the responsive profile that determines which aesthetic property is being experienced or attended to. More precisely, subject A attends to (or experiences) aesthetic property P of an object, event or scene X, when

1. A perceptually attends to X,
2. in a specific aesthetic way (as object-oriented and/or for its own sake), and
3. with a specific P-determining responsive profile.

In the next sections, and for ease of exposition, I will endorse Levinson’s own disjunctive approach to the responsive profile at play. But the account proposed doesn’t hinge on this endorsement. The wireframe definition is compatible with claims that the responsive profile has only a minimal evaluative character (e.g., DeClercq, 2002; Iseminger, 2006). As long as it is accepted that there must be some responsiveness from the subject, which goes on to fix the character of the aesthetic property being experienced, the wireframe definition is enough. In section 5, I will discuss how non-structural approaches fit into this framework and the account of joint aesthetic attention based on it.

2.2. From joint perceptual attention to joint aesthetic attention

Building on this wireframe structural account of aesthetic experience, I propose that joint aesthetic attention has a multiple-level profile. Joint aesthetic attention includes joint perceptual attention and adds two further aesthetic-specific conditions. Co-attenders jointly share their perceptual attention to the object, and also the specific aesthetic way in which they attend to it (as object-oriented, for its own sake, etc.), and their specific responsive profile (their hedonic, affective, or evaluative responses).

To better unpack this proposal, I start with joint *perceptual* attention. In contrast to parallel attention, where two individuals attend to the same object completely unaware of each other, the state of joint attention is said to be public or “out in the open” among the co-attenders. One typical way to understand the notion of “openness” is by stipulating that joint attenders enjoy some sort of “mutual awareness” of their situation (Eilan, 2005). I will later address the

question of how we should understand this condition. At this point, we can already give a preliminary definition of joint perceptual attention³:

A and B are jointly attending only when

1. A and B
 - a. perceptually attend to the same object(s) or feature(s)
and
2. A and B are mutually aware that
 - a. they are perceptually attending to such object(s) or feature(s)

Using the wireframe structure of aesthetic experience, we can thus define partial and full aesthetic attention more specifically, in a two-tier approach.

Partial joint aesthetic attention (attitude): A and B are in partial joint aesthetic attention only when

1. A and B
 - a. Perceptually attend to the same object(s) or feature(s), and
 - b. in a specific aesthetic way (object-oriented, for its own sake, etc.),
and
2. A and B are mutually aware that
 - a. they are perceptually attending to such object(s) or feature(s),
 - b. in this specific way.

Partial joint aesthetic attention merely requires that we jointly attend to the artwork or scene in an aesthetic way, one that is, for example, disinterested, for-it's-own-sake, object-focused and property-distributed, or “accompanied with understanding”. Partial joint aesthetic attention is a common phenomenon and very often a preliminarily temporal step for the full version. You attend a chamber music concert with your friend and sit silently next to each other during the performance. At the exit, you look at them enquiringly and ask, “So?”, expecting to hear their aesthetic impressions.⁴ Here, you have as yet no inkling about your friend’s response to the concert, so jointly converging onto the same aesthetic properties is as yet out of the picture. But you wouldn’t ask for their impressions if you didn’t assume that you both were engaged towards the same performance, you both attended to it aesthetically, and you are mutually aware of this, i.e., you engaged in partial joint aesthetic attention towards the performance. Now for the next tier:

³ This preliminary definition is not meant to constitute a full analysis or account of joint perceptual attention, but merely formalise the intuition that it involves some sort of mutual awareness, which separates it from other phenomena such as parallel attention.

⁴ I owe this example to a social media post by Giulia Martina.

Full joint aesthetic attention (properties): A and B are in full aesthetic attention towards aesthetic property P of an object or its perceptual features, only when

1. A and B
 - a. Perceptually attend to the same object(s) or feature(s),
 - b. in a specific aesthetic way (as object-oriented and/or for its own sake), and
 - c. with the same specific P-determining responsive profile (hedonic, affective, evaluative),and
2. A and B are mutually aware that
 - a. they are perceptually attending to such object(s) or feature(s),
 - b. in this specific way, and
 - c. with this specific responsive profile.

Armed with this definition of joint aesthetic attention, we can still bring into question: how is it possible for two agents to be mutually aware that they are attending to the same target, in the same particular (object-oriented, for-its-own-sake, etc.) way, and with the same particular responsive profile? If many aesthetic properties are constitutively dependent on an individual's response, then how is it possible to socially coordinate to them? I first return to joint *perceptual* attention in order to lay the groundwork for addressing these questions.

3. How is joint perceptual attention possible?

Any account of joint attention must specify what precisely mutual awareness is, and how it is realised by psychologically limited beings like us: how do we become mutually aware that we are attending towards the same target? There are various accounts of joint attention in the philosophical and psychological literatures. Some provide a reductive definition of mutual awareness in terms of individual mental states (e.g., Peacocke, 2005; Stueber, 2011), sometimes assumed to require recursive iterative mental representations between co-attenders (e.g., Tomasello, 2019). However, the requirement of iterative representations has been correctly criticised as psychologically implausible, even for a few steps up the inferential chain (Léon, 2021; cf. Liddle & Nettle, 2006). This criticism becomes more pointed for an iterative approach to joint *aesthetic* attention, as it may require implausible recursive inferences regarding the attentional mode and responsive profile of each other.

Other accounts argue that mutual awareness in joint attention is a relational state irreducible to the individual's states (e.g., Campbell, 2005, 2011; Seemann, 2011; Wilby, 2023; Vincini, 2024). On these views, joint attention is a primitive relational state that co-attenders share. Most of these views are concerned exclusively with joint attention as a *perceptual* phenomenon, where the target is directly and publicly available to the senses of both co-attenders (Campbell, 2018). However, under the wireframe definition I use here, aesthetic experiences require attitudes and responses that go beyond mere sensory perception. Aesthetic properties, as we have seen, are constitutively tied to the responsive profiles of perceivers. It is unclear how relational views can address the coordination of those responses.

Here, I sketch an account of joint attention that can fruitfully allow for coordinative extensions beyond purely perceptual targets, and can thus help account for joint aesthetic attention. To do so, I wish to slightly reorient the explanatory locus of joint attention. Many philosophical accounts of joint attention follow the methodological idea that we can explain mutual awareness and joint attention by specifying its truth or correctness conditions: the sufficient and necessary conditions (be they psychological, relational, spatial, phenomenological, etc.) for a situation of joint attention to be a true and justified state of the world. They give an account of the conditions under which two people are truthfully and justifiably in a state of joint attention, as preliminarily defined in the previous section (cf. Battich, 2025). At the same time, engaging in joint attention is a psychological *capacity* and, as we have seen, any analysis will be constrained by psychological plausibility. For example, 12–18-months infants are already capable of exercising this ability and, presumably, they do it not always accurately or by being sensitive to all the conditions that make joint attention a true state of the world. While being expert mindreaders, adults often don't aim for true estimates of other people's mental states, but we use efficient methods in order to advance our interactive goals, e.g., inference-to-the-best-explanation heuristics, egocentric projection, and stereotyping (Spaulding, 2018; McMahon & Isik, 2023).

In this article, I shift the locus of explanation for joint attention as having to do with *psychological capacities*. I propose a capacity-based account, addressing how people are able to engage in joint attention so that it satisfies the mutual awareness condition (see Battich, 2025, for a fuller exposition).⁵ On this approach, to possess a psychological capacity to register or recognise a thing or feature in the environment, one must be differentially sensitive to it (cf. Schellenberg, 2018; O'Callaghan, 2019). Engaging in joint attention therefore recruits the psychological capacity to distinguish when one is attending to a target together with someone. In other words, to engage in joint attention is to engage the capacity to predict and recognise when one is or not attending to the same object together with others. The task is to explain what this capacity consists in, in particular its togetherness aspect.

One can employ a capacity without being accurate, or even aiming for accuracy. Some recent philosophical and psychological approaches focus on the psychological conditions that allow us to engage in joint attention, acknowledging the limitations of our psychological resources and capacities. Julian Hauser, for example, argues that “The experience of joint attention comes first [in development], and what is added later is the ability to limit the experience to situations in which there is another agent who is jointly attending” (Hauser, 2025). Similarly, in Battich (2025), I argue that joint attention is often implicitly assumed given minimal contextual and interactive cues, even when such cues do not strictly and rationally justify it, and that curbing down this assumption can be demanding rather than the converse. Such a possibly fallible assumption is nevertheless sufficient to allow us to engage in further coordinated activities. As Sipošova and Carpenter (2019) note, the relevant notion of certainty is probabilistic. While never fully objectively certain, these possibly fallible probabilistic

⁵ This explanatory shift shares therefore some elements from primitivist views of joint attention, specifically the idea that mutual awareness or togetherness cannot be fully analysed by positing more intricate necessary or sufficient conditions. In contrast to some prominent primitivist accounts which seek to characterise the phenomenological nature of joint attention, however, I focus on psychological capacities at the individual level.

estimates will be sufficient for successful coordination and further actions based on joint attention (see also Dumas & Fairhurst, 2021; Mundy, 2018; Spaulding, 2016).

Following these recent developments, I consider the capacity to engage in joint attention as consisting of a sub-personal psychological process that can be described in probabilistic terms, which ultimately results in an output at the personal level, in which one takes oneself to be in a situation of attending to something together with someone. I then extend this approach to account for joint aesthetic attention (section 4).

Importantly, this capacity-based approach still allows me to posit the conditions for a state of joint attention to be a true state of the world. Such conditions will establish a direct connection with the definitions of joint (aesthetic) attention in section 2. On this view, it is true that A and B jointly attend to X *iff*

1. A attends to X.
2. B attends to X.
3. A takes herself to be attending to X together with B.
4. B takes herself to be attending to X together with A.

One might worry that this account is circular: conditions (3) and (4) seem to embed joint attention in its own analysis, since taking oneself to be attending together with someone just is a form of joint attention. Recall that conditions (1 – 4) merely enumerate the conditions for an instance of joint attention to be true instead of not true in the world. By themselves, they are not meant to constitute a deep or complete *explanation*. Conditions (3) and (4) refer to a psychological state, an *assumption*, that is the output of employing the capacity to predict when one is or not attending to the same object together with someone. This assumption is probabilistic, more or less right, and often fallible. The circularity worry would arise if we hold that the psychological capacity invoked in (3) and (4) itself presupposes an understanding of the truth conditions for joint attention in this account (1 – 4), indeed threatening a circular regress. But this understanding need not be conceptual, linguistic, or metaphysically accurate, and it doesn't require representing the necessary and sufficient conditions that make a situation a true joint attention situation in the world⁶. Nevertheless, there still should be *some* satisfaction criteria or understanding regarding what counts as togetherness, constituting a subjective and implicit internal threshold, that an agent uses in order to arrive at the assumptions in (3) and (4).

A key feature of a joint attention situation is that it allows for other actions and behaviours that would not be successful otherwise. Here, I propose that, in order to recognise a situation of joint attention (or to assume that one is in such a situation), a creature would require, minimally, to recognise that this situation allows for actions that would not be possible otherwise (cf. Phillips, 2021). That is, what is needed is a minimal functional sensitivity to some of the behaviours and upshots that a joint attention situation affords: that they allow for further coordinated actions and behaviours that would not otherwise be possible, that they can give rise to shared beliefs and responses, that they differ from some other situations (like

⁶ See Harder (2024), for a discussion of the circularity issue in joint activities. On Harder's view, co-attenders are not required to possess a *concept* of joint attention (or the conditions under which it is true and justified) in order to engage in it, but rather a *nonconceptual* personal-level state.

parallel or solitary attention) in ways that are behaviourally and practically significant. An agent can have *some* of this implicit functional understanding, and employ the capacity for joint attention on the basis of it, without having any explicit or reflective grasp of what joint attention, mutual awareness or togetherness really are.

Take this extreme case (modified from Phillips, 2021): a deluded person believes that everyone is constantly in situations of joint attention towards everything she attends to (e.g., when conversing on the phone, she refers to close-by objects as if the absent interlocutor were physically there too, assuming they are jointly attending with her to these objects). The deluded person has some inaccurate understanding of joint attention situations, but also some important accurate understanding. She still understands that a situation of joint attention allows for some further joint activities regarding the object, like conversing about it. On a capacity-based approach, this is sufficient for the person to qualify as having *an* understanding of joint attention. And such understanding provides sufficient internal implicit criteria to support a capacity to recognise (what this person takes to be) situations of joint attention from other situations. Understanding which situations afford which further actions and behaviours comes in degrees, and it is learned through development (cf. Hauser, 2025). This account of joint attention relaxes some of our intuitive requirements for joint attention, since it allows that, given the right conditions (1 - 4) hold, even our deluded person will be in a true joint attention situation some of the time. This relaxation is, in part, what makes the approach explanatorily useful and readily extensible in order to account for joint aesthetic attention. I will address this point in more detail in the next section.

The employment of this capacity can be further specified. Based on recent psychological work (Mundy, 2018; Siposova & Carpenter, 2019; Koster-Hale & Saxe, 2013), the employment of the capacity to engage in joint attention can be specified as involving a decision problem facing three kinds of uncertainties or probabilities:

1. Degrees of certainty about the object of my perceptual attention.
2. Degrees of certainty that the other attends to the same object as me.
3. Degrees of certainty that the other's attention is being made available to me.

It is worthwhile to distinguish between the problem a psychological system must solve, and the way it solves it.⁷ Therefore, by these probabilities, I refer to uncertainties that the psychological system faces, without suggesting that they have to be directly represented as such by the individual. There are several ways in which a system can estimate these probabilities (or face these uncertainties). More likely, this will involve mutually inclusive subpersonal processes of learning, expectations, and heuristic assumptions of different kinds (cf. Vincini, 2023; 2024; Reddy, 2010).

Estimating the first kind of probability requires, for instance, the processing of multiple sense modalities and sensory noise (e.g., "Can I both see and hear the target object, or merely hear it? Do I hear it well?"). Here, I take it that for most situations of joint attention, the degree of certainty is high: I am fairly certain of where my perceptual attention is focused. Estimating the second probability will in turn require gathering information about the other's attentional

⁷ The probabilities listed are attributable states of the system even if their estimation doesn't occur through mechanisms that explicitly represent them as such (see, e.g., Piccinini, 2015; Egan, 2017).

state, and about the target of joint attention. The third probability concerns something more specific: not merely where the other's attention is directed, but whether they are making their attentional state outwardly available in a way that I can detect and respond to. This will usually require some interactive cues, or at least the assumption of interaction⁸. Gathering this information and the estimations based on it, plausibly involves the low-effort application of heuristics and assumptions at different levels. This information and how we access it are not merely visual, as attention coordination is supported by multiple sensory modalities (Battich et al., 2020). For example, while roaming the art gallery I hear, but do not directly see, that you are nearby, so I just assume that you are next to me, and therefore I also assume that you perceive the painting directly in front of us.

Merely estimating these uncertainties is not sufficient, however, to assume one is in a joint attention situation. Following current psychological theories of perceptual decision making under uncertainty (Macmillan & Creelman, 2005; Ratcliff et al, 2016), the output will depend on whether the estimations cross a subjective threshold, referring to the implicit criteria and understanding which that particular subject uses to recognise and differentiate attentional situations. Whenever the estimations cross this threshold, the individual then takes themselves to be in a situation where they are attending to the same target together with someone, a situation which affords further actions regarding that target and that someone. At a phenomenological level, this may result in what is usually understood as a “sense of us” or of “we” (Gallagher, 2010; Eilan, 2015), although I make not commitments here on this point.⁹

It is not necessary for agents to arrive at a high degree of certainty in these estimations. On the contrary, given some interactive cues and background assumptions, a minimal step would be to implicitly *assume* that we already *are* attending together to the same target. This picture is further supported by recent developmental evidence (see Battich, 2025, for more details). Hauser (2025) argues that infants begin by tacitly assuming that all interactions are open, and the capacity for joint attention emerges not when they first experience openness, but when they develop sufficient sensitivity to detect and exclude non-open interactions from that default assumption. This picture agrees as well with the recently proposed “altercentric hypothesis” of infant development (Southgate, 2020; Grosse Wiesmann & Southgate, 2021), which holds that infants are inherently other-centred. Rather than struggling to overcome their own perspective in order to access that of others, infants preferentially encode the world from others' perspectives from very early in life.

One advantage of the capacity-based account is that it can easily be extended to account for how we coordinate not just our perceptual attention, but non-perceptual states as well.

⁸ I say *usually*, because I could successfully estimate whether the other's attention is made available to me without any overt interaction, based instead solely on background contextual cues and expectations. If we are watching a tennis match from the stands next to each other, I can expect (and assume) that you are attending to the match and that your attention to it is publicly available to others, including me, given the usually reasons, expectations and background assumptions about people that go to tennis matches (cf. Fernández-Castro & Pacherie, 2023).

⁹ In contrast to Campbell (2005; 2011), Gallagher (2010), Wilby (2023), Vincini (2024) and others, I do not aim here to characterise the phenomenological aspects of joint attention. The focus of this proposal is on the underlying psychological capacities and processes. For this reason, it may be in principle compatible with some aspects of these phenomenology-based accounts.

4. How is joint aesthetic attention possible?

4.1. From joint perceptual attention to joint aesthetic attention

This approach can be naturally developed to give an account of how joint aesthetic attention is possible. The challenge for two agents is to solve a social coordination problem towards the same aesthetic property. Establishing (full) joint aesthetic attention can then be seen as involving the estimation of several probabilities having to do with:

1. Degrees of certainty about
 - a. the object of my perceptual attention,
 - b. the specific way I attend to it, and
 - c. my specific (hedonic/affective/evaluative) response to it.
2. Degrees of certainty that the other attends to
 - a. the same object as me,
 - b. in the same specific way, and
 - c. with the same responsive profile.
3. Degrees of certainty that the other makes available to me
 - a. their perceptual attention,
 - b. the way in which they attend to the object, and
 - c. the specific responsive profile.

(For partial joint aesthetic attention, remove subconditions (c) from each probability). Estimating all these probabilities may seem a daunting task and rather cognitively demanding. Here, again, the problem that a psychological system must solve should be distinguished from the way it solves it. The three nested probabilities are merely a description of the probabilistic problem or uncertainty challenge that we face in order to engage in joint aesthetic attention, at a suitably high abstract level of description.

As with joint perceptual attention, facing up to this challenge will likely involve subpersonal processes of learning, predicting, and various heuristic assumptions. At a minimum, it will require gathering information about the other's attentional state, and not merely *what* they are attending to, but also their *way* of attending to it, and their specific responsive profile. Such gathering requires, in turn, different levels of assumptions, which likely will involve heuristic mechanisms, expectations, and previously learned background knowledge. I will address which requirements may be involved in such assumptions in the next subsection.

Again, an individual doesn't need to estimate these probabilities to any high degree of certainty, as long as they cross a subjective implicit threshold. Given some interactive cues and background assumptions, a minimal step would be to *implicitly assume*, in a heuristic-like manner, that we already *are* attending together to the same aesthetic properties or having the same aesthetic experience. To engage in joint aesthetic attention therefore consists of engaging

sub-personal psychological processes that can be described in probabilistic terms, which ultimately result in an output at the personal level: I implicitly assume that we are in a situation of joint aesthetic attention, a situation which affords further actions and behaviours regarding the object, the way we attend to it, and the aesthetic properties we experience.

A large variety of behavioural cues and information can be readily harnessed to estimate whether we share the same attentional for-its-own-sake profile and the same responsive profile toward the same target. These can include bodily expressions and gesture, pre-reflective mirroring and synchronisation (Bekkering et al., 2009; Dumas & Fairhurst, 2021; Rizzolatti & Fogassi, 2014), vocal inflection, reciprocal cues and feedback, as well as cultural knowledge and expectations. Polite (2019) notes that “part of our choice to, say, attend a concert of our favourite band together is the implicit (and reasonable) *expectation* that you and I will respond similarly to songs we both already enjoy, which we have listened to and sung together countless times” (2019, p. 441, my emphasis). Even the decision to visit the art gallery together provides a background of cultural norms and expectations regarding our responses to artworks. Many of our shared cultural norms come with coordinative commitments, which “enable agents to have more reliable expectations about their partners’ actions than would otherwise be possible, thus reducing uncertainty and facilitating cooperation and coordination” (Fernández-Castro & Pacherie, 2023, p. 893), and providing reasons to engage in joint activities. From a normative approach, joint attention itself may be understood as involving shared commitments to attend to the same target (Scarafone, 2024). The converse of commitments are entitlements: for example, given our cultural norms around attending chamber music concerts, I can be entitled to assume that you engage towards the concert aesthetically, even when you don’t (cf. Geurts, 2019).

It is now possible to spell out an external account of joint aesthetic attention, specifying the conditions under which a state of joint aesthetic attention is a true state of the world. Such conditions connect the capacities-based probabilistic approach I offer here to the definition of joint aesthetic attention in section 2. It is true that A and B jointly aesthetically attend to aesthetic property P of object(s) or feature(s) X *iff*

1. A
 - a. Perceptually attends to X,
 - b. in a specific aesthetic way (as object-oriented, for its own sake, etc.),
 - c. with a specific P-determining responsive profile R (hedonic, affective, evaluative, etc.).
2. B
 - a. Perceptually attends to X,
 - b. in a specific aesthetic way,
 - c. with specific P-determining responsive profile R.
3. A takes herself to be
 - a. perceptually attending to X together with B,

- b. in a specific aesthetic way together with B,
 - c. with specific responsive profile R together with B.
- 4. B takes herself to be
 - a. perceptually attending to X together with A,
 - b. in a specific aesthetic way together with A,
 - c. with specific responsive profile R together with A.

This account of the external true conditions for joint aesthetic attention can successfully elucidate when it differs from other situations. In a case of parallel aesthetic attention, (1) and (2) are true, but (3) and (4) are not. However, we can also have failures at various different levels. Suppose a situation where (2.c) and (4.c) is false: when admiring an artwork, I take it that we both see it together as beautiful (I assume a situation of full joint aesthetic attention), while you take it that we are both admiring it aesthetically (you assume a situation of partial joint aesthetic attention), as you experience it as grotesque.

Conditions (3) and (4) refer to the psychological *assumptions* that A and B each adopt when facing up the probabilistic uncertainty challenge of joint aesthetic attention. More simply put, an instance of joint aesthetic attention is a true instance in the world only when A and B each experience the same aesthetic property of an object, event or scene, and each of them assumes that they are doing it jointly. Of course, as the probabilistic account above shows, these assumptions do not occur in a vacuum, but rely on implicit heuristic assumptions, expectations, and the dynamic gathering and sharing of social cues and background information (cf. Gallotti et al., 2017). Establishing and maintaining joint aesthetic attention can and very often is a dynamic process. There is usually a coordinated to-and-fro between two co-attenders, mediated by behavioural cues and implicit, potentially fallible, assumptions and estimates regarding these cues, which in turn can lead to further behaviours and assumptions from each subject in such a way that they feed off each other.

Nevertheless, this account of joint aesthetic attention, and that of joint attention in general on which it rests, does impose weaker requirements than other accounts, which often focus on explaining what gives two people immediate and appropriate justification to take their attention as being joint (e.g., Peacocke, 2005; Campbell, 2005; Seemann, 2024; Harder, 2024; among many others). As a consequence, on this account more situations count as true joint aesthetic attention than our intuitions might allow. This is mainly due to two features.

First, this account separates the truth conditions for a joint aesthetic attention situation to occur in the world, from the conditions that may justify each subject's being in such a state. On the present account, all that is required in conditions (3) and (4) is that each subject engage in a possibly fallible estimation process and arrive at the assumption that they are in a situation of joint aesthetic attention. It does not require, however, that such assumptions be properly externally justified.¹⁰ They need not be factive. Second, employing the capacity to enter into joint attention situations (both perceptual and aesthetic) requires only a minimal

¹⁰ Taking a different approach, Harder (2024) appeals to the notion of practical justification based on appropriate motivational states, which is in some respects weaker than theoretical justification based on truth conditions. However, my approach here is weaker still.

functional understanding of some of the behaviours and upshots that joint attention situations afford. For this reason, my account considers as true joint aesthetic attention some situations where it seems, intuitively, that subjects got there through mere luck. The most glaring case would be when people have a non-typical implicit understanding to recognise joint aesthetic attention situations.

Suppose our deluded person from above has a partner: Two delusional people both believe that everyone they know is constantly in a state of joint aesthetic attention with them. They still understand that a situation of joint aesthetic attention allows for some further joint activities regarding the target, like conversing about it, that wouldn't otherwise be possible, but they assume they are constantly in such a state with all their acquaintances. Suppose they know each other. Suppose further that they happen to be in the same room of an art gallery, looking at and admiring the same sculpture, which evokes the same response in each of them. Yet they are standing on opposite sides of the sculpture and have not seen each other in the gallery yet, so they cannot be accurately aware of each other's presence, even less their respective responses.¹¹ Here, all truth conditions from (1.a, 1.b, 1.c.) to (4.a, 4.b, 4.c) hold, so this is a situation of full joint aesthetic attention on my account. But it may seem unintuitive to claim that these two people are correctly jointly attending to the same aesthetic property of the sculpture. Intuitively, they *should* be accurately aware of each other's presence and responses in order for this to count as true joint aesthetic attention.

This case might seem like an objection if we hold onto that intuition. However, the case merely highlights that the emphasis of this account is not on the conditions that justify or make true each co-attender's assumption, but on what joint attention allows us to do, its functions, and how each co-attender implicitly understands them. "Lucky" true situations of aesthetic joint attention still play this role. Suppose that these two people later do meet face to face in the vestibule, and one of them says "It was nice sharing that experience of the sculpture with you". The hearer may very well understand what the speaker is referring to. In short, they could successfully perform (granted, only some) further interactions on the basis of their joint aesthetic attention to the sculpture. As long as conditions (1 - 4) hold, these two people's minimal, incomplete understanding of which situations are joint attention situations and which aren't, does not impede them from engaging in "lucky" true joint attention situations, benefiting from many of the functions that joint attention has in our social lives, including epistemic and behavioural upshots.

4.2. Psychological preconditions for joint aesthetic attention

So far, I have offered an account of how we engage joint aesthetic attention as a social coordination problem with a multiple coordinative profile. This involves making assumptions that we are attending together to the same aesthetic properties or having the same aesthetic experience, even if they are fallible. However, we can further ask what the psychological capacities and understanding necessary to make such assumptions are. What is the implicit functional understanding needed in order to employ the capacity for joint aesthetic attention?

Regarding joint *perceptual* attention, the psychological capacity to understand what it is for another person to attend or engage towards an object, seems to be a necessary precondition in order to engage in joint attention. Since joint attention is already present in 18-month-old

¹¹ I am indebted to a reviewer of this journal for this example.

infants, this understanding is likely non-conceptual, merely involving a capacity to recognise when someone is perceptually engaging with an object, by being sensitive, for example, to the causal relations between gaze direction and other behaviours and states of the agent, such as emotions or goals (e.g., you are afraid because you are *seeing* the dog in front of you) (Butterfill & Apperly, 2013; Phillips, 2021; Wellman, 2014). This minimal understanding of “seeing” or “attending” should be sufficient to support perceptual joint attention.

Analogously, partial joint *aesthetic* attention (same aesthetic attitudes) will in turn require some minimal capacity to recognise when someone else’s attention is aesthetic, that is, when someone is engaging towards an object in an aesthetic way or with an aesthetic attitude. This capacity could of course involve the deployment of sophisticated and even very specific concepts of what the aesthetic attitude is (e.g., when a philosopher of art is engaging aesthetically towards an artwork). But, as with concepts of perceptual attention, such sophisticated conceptual involvement may not be necessary. It doesn’t seem to be necessary to have a conceptual, reflective, and well-articulated idea of what aesthetic attitude is in order to recognise when someone else is engaging in it. Joint aesthetic attenders could only have an implicit, perhaps even nebulous and underdefined, idea of aesthetic attention or attitude and still be able to recognise both when they themselves and when others are engaging in it. I could, for example, recognise that someone is engaging towards an object in an *appreciative* attitude, or in a way that makes them feel emotions that I can feel in similar situations. I don’t necessarily have to be able to name or define this way of appreciating, as this recognition is very likely embodied and effortless: *I feel* that I am engaging towards the object in a certain mode, and I can similarly feel when someone else is engaging in a similar mode.

Beyond the partial version, the necessary psychological capacities to engage in full joint aesthetic attention (same aesthetic properties) may well vary depending on the particular aesthetic quality. In order to mutually assume that we are attending together to the same aesthetic property, we should be able to at least recognise when oneself and others are experiencing said property, together with the usual cues and behaviours that such experiences elicit. Without explicit linguistic communication, we likely find it easier to coordinate towards large-grain aesthetic properties, starting with raw valence properties like “good” and “bad”, and then with other possibly salient properties like “beautiful”, “pleasant”, “sad”, “happy”, “disgusting”, “scary”, etc. Again, it is possible that for many of these coarse-grained aesthetic properties, it is not necessary for individuals to have a clear, well-defined and consciously articulated concept of it in order to recognise when they themselves and others are experiencing it, or something like it.

For more complex aesthetic properties, a deeper mastery of their cultural usage seems necessary for joint aesthetic attention, a usage that seems more explicitly about *concepts* than properties (Sibley, 1959). It is undeniable that the use of many aesthetic predicates is much closer to concept deployment than to perception of simple secondary qualities like colour. When I say that “this sculpture is tragic”, even if I intend the predicate as a property attributed to the object, a fairly intuitive way of describing what I do is that I am deploying the concept of “tragic”, a concept I have learned to use through immersion and education in Western cultural norms. The account of joint aesthetic attention developed here does allow for the joint deployment of the same aesthetic concept towards the same object, event, or scene. The responsive profile of aesthetic experience can easily include the deployment of aesthetic concepts: to experience an aesthetic property requires mastery of its corresponding aesthetic concept. Here, explicit linguistic communication is perhaps the easiest and most efficient way

to make my response overt and establish joint aesthetic attention. But even for complex aesthetic properties, there can be cases where we engage in joint aesthetic attention without linguistic communication. These are cases where the background conditions and common ground between us might facilitate, even force, the assumption that we both share the same aesthetic experience. For example, if just before attending a concert of Louise Farrenc's *Symphonie n°3* we watch together a video highlighting that her piece presents a "clear and rigorous" structural line while keeping a "persuasive, dramatic" edge, we might then take ourselves to jointly experience these qualities when we hear the concert.

Finally, what is being shared may of course not correspond to the fullness of what each of us experiences. That is, when we engage in joint attention to the aesthetic properties of an artwork, I might also be sensitive to aesthetic properties that I do not share with you. Even when both of us jointly attend to the beauty of an artwork, the "beauty" that I experience might not be exactly the same "beauty" that you experience. It is possible that each of us assumes, correctly, that we are both experiencing the beauty of the artwork together, while at the same time you feel that your own experience of this beauty is more nuanced and sophisticated than mine (say, because you are an inveterate aesthete, while I'm a novice to art appreciation). This difference doesn't prevent us from being able to successfully converge and coordinate on the same aesthetic property, being aware of what we are sharing about it, and what we might not, and using that coordination as the basis for further interaction.

These differences between what we experience together and what we experience individually also suggest that the mere attempt to engage in joint aesthetic attention, even when unsuccessful, can have a crucial function: to facilitate learning and training of members of the community on how to recognise aesthetic properties, how to evaluate artworks, scenes and events aesthetically, and in coming to understand each other's aesthetic values (Murdoch, 1970). Understanding how other people experience the aesthetic properties of a shared attentional target allows us to understand their "beliefs, concerns, and values that structure the co-attender's use of her aesthetic concepts" (Gomes, 2013, p. 332), beliefs and values that are equally at play in their moral framework.

5. Alternative frameworks

5.1. Content-oriented accounts of aesthetic experience

The main competitor to the structural family of theories on aesthetic experience is the "content" family (Nanay, 2018). According to content-oriented views, aesthetic experience is defined in terms of its content, that is, the aesthetic properties it is directed at. For these views, aesthetic properties are properties of the object, independent from the observer's attitude or response. The most advanced content-oriented views adopt what could be termed moderate formalism: only some aesthetic properties are formal in the sense of being accessible by direct sensation alone, whereas many other non-perceptual aesthetic properties supervene on or are determined by these perceptual properties (Carroll, 2006; Hick, 2012; Zangwill, 2001¹²). Thus, for Noël Carroll, one is having an aesthetic experience of an artwork just in case one is "attending with understanding to the formal and/or the aesthetic and/or the expressive properties of a work of art" (2012, p. 173; see also Carroll, 2000, 2015). What unifies all these properties under the aegis of the aesthetic (in the wider sense of our enquiry, not Carroll's

¹² Admittedly, Zangwill's view is not properly a content view (see subsection 5.2).

narrow sense) is that such properties are the “ensemble of choices elected to realise the point or the purpose of the artwork” (Carroll, 2012, p. 173).

Although Carroll doesn’t endorse a response-dependence view of aesthetic properties, he does highlight an important subjective requirement for experiencing them: attention “with understanding”. What puts the aesthetic in aesthetic experiences is the content, but in order to experience such content at all, you need to attend to it, and do so in a particular way. I will not attempt an exegesis of Carroll’s “attention with understanding” (cf. Vernazzani, 2023). The key point is that, even for a content-view that doesn’t endorse response-dependence, there is an important subjective element involved in aesthetic experiences.

Interestingly, the first two clauses of my wireframe definition of aesthetic experience can be made to encompass Carroll’s content view: Aesthetic experience is experience involving (1) perception of some object, event or scene, (2) grounded in aesthetic attention towards that object, event or scene. One can fill in the details about aesthetic attention in different ways, Carroll’s “attention with understanding” being one of many. Granted, Carroll will enlarge the first clause to involve a particular kind of content, but the key point is that experiencing such content still necessitates a subjective element. The wireframe definition makes good his idea: aesthetic experience is grounded in a particular sort of aesthetic attention towards the object, event or scene.

How precisely Carroll’s view fits within my account of partial and full joint aesthetic attention will depend on how his theory determines *which* aesthetic property we attend to. Carroll is not explicit on this issue, so here I attempt my best interpretation. On the content-oriented view, when I apply the right sort of “attention with understanding”, I will experience the right aesthetic content. If so, then, as long as we engage jointly with the same kind of attention, we will both jointly experience the same aesthetic content. On this interpretation of Carroll’s view, full joint aesthetic attention is not only possible but easily explained. What requires care, then, is to explain how we can engage in partial joint aesthetic attention: when we jointly attend to the artwork with an aesthetic attitude, but fail to experience the same aesthetic properties. I suspect that here the object-oriented supporter will have to give us more details of what “attention with understanding” is, and how it could support the fairly common phenomenon of partial joint aesthetic attention.

5.2. Anti-realism and subjectivism about aesthetic properties

Nick Zangwill (2012) has notoriously argued that we may hear the physical sounds of music together, but we cannot attend together to the same aesthetic properties of the music. We remain in “aesthetic solitary confinement” (2012, p. 384). This conclusion should follow, so he argues, from his metaphysical commitments about aesthetic experience and aesthetic properties (see Polite, 2019 for further discussion).

Zangwill holds that sensory formal properties partly determine other (non-sensory) aesthetic properties, together with “any dispositions to provoke responses that might be thought to be partly constitutive of aesthetic properties” (Zangwill, 2001, p. 57). This response-dependent character leads him to endorse an anti-realist position regarding aesthetic properties, where aesthetic properties are not part of the “real world”. It also leads him into a subjectivist position: many aesthetic qualities are not ontological properties of the object, but of the subject’s dispositions (see, e.g., parts 1 and 3 in Zangwill, 2001). Zangwill’s dispositional response-dependent approach has led him to conclude that “there is no joint perception of

aesthetic properties” (Zangwill, 2012). More generally, joint aesthetic attention does seem intuitively difficult if we commit to anti-realism and subjectivism about aesthetic properties.

However, as Elisabeth Schellekens (2008) correctly argues, the response-dependency question can, and should, be held separately from the typical metaphysical debate about whether aesthetic properties are real or not. The more so when response-dependence can be co-opted in favour of both realist (e.g., Levinson, 2005; Pouivet, 2006; cf. Pettit, 1991; Wright, 1992) and anti-realist (e.g., Matravers, 2005; Scruton, 2010; Zangwill, 2001) positions. Schellekens (2008) also identifies as a separate question the issue of whether aesthetic properties are to be ascribed to the object (objectivism) or the subject (subjectivism). As with the real/anti-real debate, this issue is orthogonal to the response-dependency question, as response-dependence is compatible with both objectivist (e.g., Levinson, 2006) and subjectivist (e.g., Scruton, 2010; Zangwill, 2001) positions.

My wireframe definition of aesthetic experience is not committed to any particular metaphysical position about realism or subjectivism/objectivism, merely to some form of response-dependency. This definition, and the account of joint attention built on it, can then easily accommodate Zangwill’s dispositional response-dependent approach. As long as some response-dependency thesis is at play, both anti-realism and subjectivism allow for the possibility that two observers could have the same aesthetic experience towards the same target, and be aware that they both do so. The positive account of how we come to engage in joint aesthetic attention I develop here should, I hope, help sway even staunch anti-realists and subjectivists about aesthetic properties that we are not in an “aesthetic solitary confinement.”

6. Conclusion

I have introduced and analysed the notion of joint aesthetic attention as a social coordination phenomenon with a multiple coordinative profile. Co-attenders must coordinate not only their perceptual focus on a target, but also the specific aesthetic way they attend to it, and their specific responsive profiles, be they hedonic, affective, and/or evaluative. I have argued that this multiple coordination is made possible through psychological capacities that implement a probabilistic process, which relies on the dynamic interplay of implicit assumptions, behavioural cues, shared background expectations, and reciprocal feedback, to estimate and coordinate our aesthetic experiences.

While some aspects of the proposal remain to be further detailed, the framework advanced here offers a principled starting point for further philosophical and empirical inquiry on joint aesthetic attention. One relevant open question concerns large groups. Can we have *collective* aesthetic attention, i.e., collective aesthetic experiences towards the same aesthetic properties of a target, and if so, how? It is in principle possible to extend my proposal from joint to collective aesthetic attention. But such an extension is not straightforward. Any proposal of group aesthetic experiences will have to address some further crucial questions. One important issue has to do with the scope of group awareness necessary for collective aesthetic attention. On one approach, everyone in the group has to be aware that everyone in the group shares their aesthetic attention towards the same properties. Another approach may relax the scope, holding that my assumption of collective aesthetic attention is correct merely when the majority of group members share that assumption with me. In addition, social expectations and the shared background and norms in the group may play a larger role here, overriding the

requirement that everyone or the majority be aware of sharing their aesthetic attention towards the same target.

Acknowledgements

I am grateful to Claudia Cano, Elisabeth Pacherie, Denis Buehler, Jérôme Dokic, and the reviewers for this journal, as well as audiences at the “Aesthetics, Cognition & Environment” Workshop (Paris) and the SoPhA Congress 2025 (Nantes), for their valuable comments on earlier versions of this paper.

The author acknowledges the support of grants from the French National Research Agency while based at the École Normale Supérieure (grant ANR-23-AERC-0016 awarded to Lucas Battich., and grant ANR-17-EURE-0017 FrontCog, ENS-PSL).

For the purpose of Open Access, the author has applied a CC-BY public copyright license (<https://creativecommons.org/licenses/by/4.0/>) to the present document and to all subsequent versions up to the Author Accepted Manuscript arising from this submission.

References

- Battich, L. (2025). Explaining Joint Attention: Between Epistemic Justification and Psychological Processing. *Philosophical Psychology*, 1–28. <https://doi.org/10.1080/09515089.2025.2538752>
- Battich, L., Fairhurst, M., & Deroy, O. (2020). Coordinating Attention Requires Coordinated Senses. *Psychonomic Bulletin & Review*, 27(6), 1126–1138. <https://doi.org/10.3758/s13423-020-01766-z>
- Bekkering, H., De Bruijn, E. R. A., Cuijpers, R. H., Newman-Norlund, R., Van Schie, H. T., & Meulenbroek, R. (2009). Joint Action: Neurocognitive Mechanisms Supporting Human Interaction. *Topics in Cognitive Science*, 1(2), 340–352. <https://doi.org/10.1111/j.1756-8765.2009.01023.x>
- Bruner, J. S. (1974). From Communication to Language: A Psychological Perspective. *Cognition*, 3(3), 255–287. [https://doi.org/10.1016/0010-0277\(74\)90012-2](https://doi.org/10.1016/0010-0277(74)90012-2)
- Butterfill, S. A., & Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind and Language*, 28(5).
- Campbell, J. (2005). Joint Attention and Common Knowledge. In N. M. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology* (pp. 287–297). Oxford University Press.
- Campbell, J. (2011). An Object-Dependent Perspective on Joint Attention. In A. Seemann (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience* (pp. 415–430). MIT Press.
- Campbell, J. (2018). Joint Attention. In M. Jankovic & K. Ludwig (Eds), *The Routledge Handbook of Collective Intentionality* (pp. 115–129). Routledge.
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, 63(4), i–174. <https://doi.org/10.2307/1166214>

- Carroll, N. (2000). Art and the domain of the aesthetic. *The British Journal of Aesthetics*, 40(2), 191–208. <https://doi.org/10.1093/bjaesthetics/40.2.191>
- Carroll, N. (2006). Aesthetic Experience: A Question of Content. In M. Kieran (Ed.), *Contemporary debates in aesthetics and the philosophy of art* (pp. 69–97). Blackwell.
- Carroll, N. (2012). Recent Approaches to Aesthetic Experience. *The Journal of Aesthetics and Art Criticism*, 70(2), 165–177. <https://doi.org/10.1111/j.1540-6245.2012.01509.x>
- Carroll, N. (2015). Defending the Content Approach to Aesthetic Experience. *Metaphilosophy*, 46(2), 171–188. <https://doi.org/10.1111/meta.12131>
- Cochrane, T. (2009). Joint attention to music. *British Journal of Aesthetics*, 49(1), 59–73.
- DeClercq, R. (2002). The Concept of an Aesthetic Property. *The Journal of Aesthetics and Art Criticism*, 60(2), 167–176. <https://doi.org/10.1111/1540-6245.00063>
- Deroy, O., Longin, L., & Bahrami, B. (2024). Co-perceiving: Bringing the social into perception. *WIREs Cognitive Science*, n/a(n/a), e1681. <https://doi.org/10.1002/wcs.1681>
- Dumas, G., & Fairhurst, M. T. (2021). Reciprocity and alignment: Quantifying coupling in dynamic interactions. *Royal Society Open Science*. <https://doi.org/10.1098/rsos.210138>
- Eilan, N. (2005). Joint Attention, Communication, and Mind. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds), *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology* (pp. 1–33). Oxford University Press.
- Eilan, N. (2015). *Joint Attention and the Second Person (Draft)*. <https://warwick.ac.uk/fac/soc/philosophy/people/eilan/jaspup.pdf>
- Egan, F. (2017). Function-theoretic explanation and the search for neural mechanisms. In D. M. Kaplan (Ed.), *Explanation and integration in mind and brain science* (pp. 145–163). Oxford University Press.
- Fernández-Castro, V., & Pacherie, E. (2023). Commitments and the sense of joint agency. *Mind & Language*, 38(3), 889–906. <https://doi.org/10.1111/mila.12433>
- Gallotti, M., Fairhurst, M. T., & Frith, C. D. (2017). Alignment in Social Interactions. *Consciousness and Cognition*, 48, 253–261. <https://doi.org/10.1016/j.concog.2016.12.002>
- Geurts, B. (2019). Communication as commitment sharing: Speech acts, implicatures, common ground. *Theoretical Linguistics*, 45(1-2), 1–30. <https://doi.org/10.1515/tl-2019-0001>
- Gomes, A. (2013). Iris Murdoch on Art, Ethics, and Attention. *The British Journal of Aesthetics*, 53(3), 321–337. <https://doi.org/10.1093/aesthj/ayto20>
- Harder, R. (2024). Joint attention and communication. *Inquiry*, 67(10), 3796–3834. <https://doi.org/10.1080/0020174X.2022.2074101>
- Hauser, J. (2025). Joint Attention, Openness, and Self–Other (In)Differentiation. *Journal of Consciousness Studies*, 32(1-2), 50–75. <https://doi.org/10.53765/20512201.32.1.055>
- Hick, D. H. (2012). Aesthetic Supervenience Revisited. *The British Journal of Aesthetics*, 52(3), 301–316. <https://doi.org/10.1093/aesthj/ays024>

- Hobson, P., & Hobson, J. (2011). Joint Attention or Joint Engagement? Insights from Autism. In A. Seemann (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience* (pp. 115–136). MIT Press.
- Iseminger, G. (2006). The Aesthetic State of Mind. In M. Kieran (Ed.), *Contemporary debates in aesthetics and the philosophy of art* (pp. 98–110). Blackwell.
- Johnston, M. (1989). Dispositional Theories of Value III. *Aristotelian Society Supplementary Volume*, 63(1), 139–174. <https://doi.org/10.1093/aristoteliansupp/63.1.89>
- Kampis, D., & Southgate, V. (2020). Altercentric cognition: How others influence our cognitive processing. *Trends in Cognitive Sciences*, 24(11), 945–959. <https://doi.org/10.1016/j.tics.2020.09.003>
- Koster-Hale, J., & Saxe, R. (2013). Theory of Mind: A Neural Prediction Problem. *Neuron*, 79(5), 836–848. <https://doi.org/10.1016/j.neuron.2013.08.020>
- León, F. (2021). Joint Attention without Recursive Mindreading: On the Role of Second-Person Engagement. *Philosophical Psychology*, 34(4), 550–580. <https://doi.org/10.1080/09515089.2021.1917533>
- Levinson, J. (2005). Aesthetic properties II. *Aristotelian Society Supplementary Volume*, 79(1), 211–227. <https://doi.org/10.1111/j.0309-7013.2005.00133.x>
- Levinson, J. (2006). Aesthetic Properties, Evaluative Force, and Differences of Sensibility. In J. Levinson (Ed.), *Contemplating Art: Essays in Aesthetics* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199206179.003.0020>
- Levinson, J. (2016). *Aesthetic Pursuits: Essays in Philosophy of Art*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198767213.001.0001>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd Ed. Lawrence Erlbaum Associates Publishers.
- Matravers, D. (2005). Aesthetic Properties I. *Aristotelian Society Supplementary Volume*, 79(1), 191–210. <https://doi.org/10.1111/j.0309-7013.2005.00132.x>
- McMahon, E., & Isik, L. (2023). Seeing social interactions. *Trends in Cognitive Sciences*, 27(12), 1165–1179. <https://doi.org/10.1016/j.tics.2023.09.001>
- McNally, L., & Stojanovic, I. (2017). Aesthetic Adjectives. In J. O. Young (Ed.), *The Semantics of Aesthetic Judgements* (pp. 17–37). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198714590.003.0002>
- Moll, H. (2024). What We Do and Don't Know About Joint Attention. *Topoi*, 43(2), 247–258. <https://doi.org/10.1007/s11245-023-09961-y>
- Moore, C., & Dunham, P. J. (Eds.). (1995). *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum.
- Mundy, P. (2018). A Review of Joint Attention and Social-Cognitive Brain Systems in Typical Development and Autism Spectrum Disorder. *European Journal of Neuroscience*, 47(6), 497–514. <https://doi.org/10.1111/ejn.13720>
- Murdoch, I. (1970). *The Sovereignty of Good*. Routledge.
- Nanay, B. (2014). Aesthetic Attention. *Journal of Consciousness Studies*, 22(5-6), 96–118.

- Nanay, B. (2016). *Aesthetics as Philosophy of Perception*. Oxford University Press.
- Nanay, B. (2018). Multimodal Mental Imagery. *Cortex*, 105, 125–134. <https://doi.org/10.1016/j.cortex.2017.07.006>
- O’Callaghan, C. (2019). *A Multisensory Philosophy of Perception*. Oxford University Press.
- Peacocke, C. (2005). Joint Attention: Its Nature, Reflexivity, and Relation to Common Knowledge. In N. M. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology* (pp. 298–324). Oxford University Press.
- Pettit, P. (1991). Realism and Response-Dependence. *Mind*, 100(4), 587–626. <https://www.jstor.org/stable/2255012>
- Phillips, B. (2021). Seeing Seeing. *Philosophy and Phenomenological Research*, 102(1), 24–43. <https://doi.org/10.1111/phpr.12636>
- Polite, B. (2019). Shared Musical Experiences. *The British Journal of Aesthetics*. <https://doi.org/10.1093/aesthj/ayz024>
- Pouivet, R. (2006). *Le Réalisme Esthétique*. Presses Universitaires de France.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, 20(4), 260–281. <https://doi.org/10.1016/j.tics.2016.01.007>
- Reddy, V. (2010). *How infants know minds*. Harvard University Press.
- Rizzolatti, G., & Fogassi, L. (2014). The mirror mechanism: Recent findings and perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1644), 20130420. <https://doi.org/10.1098/rstb.2013.0420>
- Roessler, J. (2005). Joint Attention and the Problem of Other Minds. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology* (pp. 230–259). Oxford University Press.
- Scarafone, A. (2024). Joint Attention: Normativity and Sensory Modalities. *Topoi*, 43(2), 283–294. <https://doi.org/10.1007/s11245-023-10004-9>
- Schellekens, E. (2008). Three Debates in Meta-Aesthetics. In K. Stock & K. Thomson-Jones (Eds.), *New Waves in Aesthetics* (pp. 170–187). Palgrave Macmillan UK. https://doi.org/10.1057/9780230227453_9
- Schellenberg, S. (2018). *The Unity of Perception*. Oxford University Press.
- Scruton, R. (2010). Public Text and Common Reader. In E. S. Shaffer (Ed.), *Comparative Criticism: Volume 4, The Language of the Arts* (pp. 85–106). Cambridge University Press.
- Seemann, A. (2011). Joint Attention: Toward a Relational Account. In A. Seemann (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience* (pp. 183–202). MIT Press.
- Seemann, A. (2024). Joint Attention as the Base of Common Knowledge and Collective Intentionality. *Topoi*, 43(2), 259–270. <https://doi.org/10.1007/s11245-024-10011-4>
- Sibley, F. (1959). Aesthetic Concepts. *The Philosophical Review*, 68(4), 421. <https://doi.org/10.2307/2182490>

- Siposova, B., & Carpenter, M. (2019). A New Look at Joint Attention and Common Knowledge. *Cognition*, 189, 260–274. <https://doi.org/10.1016/j.cognition.2019.03.019>
- Spaulding, S. (2016). Mind Misreading. *Philosophical Issues*, 26(1), 422–440. <https://doi.org/10.1111/phils.12070>
- Spaulding, S. (2018). *How We Understand Others: Philosophy and Social Cognition*. Routledge.
- Stolnitz, J. (1961). On the Origins of "Aesthetic Disinterestedness". *The Journal of Aesthetics and Art Criticism*, 20(2), 131–143. <https://doi.org/10.2307/427462>
- Stueber, K. R. (2011). Social Cognition and the Allure of the Second-Person Perspective: In Defense of Empathy and Simulation. In A. Seemann (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience* (pp. 265–292). MIT Press.
- Tomasello, M. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard University Press.
- Vernazzani, A. (2023). How Artworks Modify Our Perception of the World. *Phenomenology and the Cognitive Sciences*, 22(2), 417–438. <https://doi.org/10.1007/s11097-021-09742-1>
- Vincini, S. (2023). Taking the mystery away from shared intentionality: The straightforward view and its empirical implications. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1068404>
- Vincini, S. (2024). Joint attention, relationalism, and individuation. *Philosophical Psychology*, 38(5), 2024–2049. <https://doi.org/10.1080/09515089.2024.2318421>
- Wellman, H. M. (2014). *Making Minds: How Theory of Mind Develops*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199334919.001.0001>
- Wilby, M. (2023). The form and function of joint attention within joint action. *Philosophical Psychology*, 36(1), 134–161. <https://doi.org/10.1080/09515089.2022.2039384>
- Wright, C. (1992). *Truth and Objectivity*. Harvard University Press.
- Zangwill, N. (1992). UnKantian Notions of Disinterest. *The British Journal of Aesthetics*, 32(2), 149–152. <https://doi.org/10.1093/bjaesthetics/32.2.149>
- Zangwill, N. (2001). *The Metaphysics of Beauty*. Cornell University Press.
- Zangwill, N. (2012). Listening to Music Together. *The British Journal of Aesthetics*, 52(4), 379–389. <https://doi.org/10.1093/aesthj/ayso38>