

# The Textual, the Audiovisual, and Videographic Thought

**Ben Spatz**

*(they/he) is a nonbinary scholar-practitioner working at the intersections of artistic research and critical theories of embodiment and identity. They are the author of Race and the Forms of Knowledge: Technique, Identity, and Place in Artistic Research (2024), founding editor of the videographic Journal of Embodied Research, and a Reader in Media and Performance at University of Huddersfield.*

## Abstract

This essay draws on the author's experience as founding editor of the videographic *Journal of Embodied Research (JER)*, as well as their own artistic research practice and critical theories of embodiment and identity, to examine shifting relationships among the textual, the audiovisual, and the videographic. Addressing each term in sequence, the essay builds on the idea of embodied research, and the experience of developing a style guide for *JER*, to rethink the textual and the audiovisual in the context of the videographic. As the space of videographic thought becomes ever more fluid and all-encompassing, it is incumbent upon filmmakers of all kinds to critically reexamine the ways in which video remains entangled with bodies, places, and the still-powerful technology of the written word. To support such a reexamination, approaches to academic filmmaking and the video essay should be put in conversation with practices of embodied research.

**Keywords:** practice research; embodied knowledge; video essay; media ontology; decolonial theory

In 2017, I founded the *Journal of Embodied Research* (<https://jer.openlibhums.org>). Published by the Open Library of Humanities, *JER* is “the first peer reviewed, open access, academic journal to focus on the dissemination of embodied knowledge through video.” As of this writing, the journal has published nine issues comprising 35 video articles and six editorial video essays. Based in performing arts, but with highly interdisciplinary ambitions, many of the journal’s contributors are dancers, musicians, and theatre-makers who have not previously crafted works in the medium of video. Others are filmmakers who produce artistic works onscreen but are less familiar with the context of academic publishing. In this essay I consider what I have learned about videographic form, both as editor of *JER* and in ongoing conversation with the field of videographic film criticism.

In an earlier essay, published alongside the founding of *JER*, I discussed what I then called “the video way of thinking” (Spatz 2018). While that essay equates “video” with the audiovisual, this one focuses on the entangled yet still important differentiations between the textual, the audiovisual, and the videographic as modes of thought. I am a relative newcomer to film and media studies, hence the references and conceptual background of this essay may be less familiar to some readers. Yet for nearly a decade I have been working to bring experimental performing arts into richer contact with practices of video recording, editing, and publication. My most recent video essay interrogates the corporeality of whiteness in conversation with the ontological turn in recent black and indigenous studies (Spatz et al. 2022). My most recent monograph proposes a decolonial media ontology, critically analysing the differential racialization and coloniality of dominant and emerging forms of knowledge (Spatz 2024). This essay attempts to articulate what I have learned from these projects in a new disciplinary context, with the aim of supporting further conversation.

### On the Textual

As a scholarly journal publishing video, *Journal of Embodied Research* (*JER*) joins a growing landscape of audiovisual and videographic journals that includes not only *Screenworks* and *[in]Transition* but also the nonlinear, exposition-based *Journal for Artistic Research* and others that make use of its Research Catalogue platform. Within

that landscape, *JER*'s specificity is found in both its foregrounding of embodiment — understood at least initially as a “first affordance” that precedes all mediated forms (Spatz 2017) — and its approach to the form of the scholarly journal article. In launching *JER*, my intention was not primarily to put existing video works into the academic domain, but rather to investigate and experiment with the institutional form of the scholarly journal. For this reason, *JER* has always defined the “video article” in a way that is simultaneously narrow and capacious. A video article, for *JER*, is a video document or file that has been through a process of peer review. This definition is narrow in the sense that it excludes many of the multimedia forms developed by other journals (such as a video accompanied by a research statement, or a nonlinear composition comprising audio, video, image, and textual elements), as well as nonlinear forms of video (such as 360° video, or a work comprising multiple parallel video channels). Yet the definition is very broad in its approach to content, which is defined solely in relation to the peer review process and not by reference to any particular method or criterion, such as the need to make an argument or answer a question.

Perhaps the most significant point in the above definition of a video article is that *JER* does not publish any written material apart from what appears (whether spoken or written) in the video. Even the core metadata that is required for an article to exist within scholarly publishing infrastructures (title, authors, keywords, abstract) must appear first and foremost in the video document itself, which is always considered the version of record. The absence of an accompanying research statement or other writing means that the question of textuality for *JER* is always posed within, not alongside or external to, the question of videographic thought. In other words, textuality for *JER* is something that appears first of all within the space of video and can then secondarily be transcribed. Since its inception, *JER* has always published a transcript of each of its video articles. These PDF and XML documents are not the article itself, but merely transcripts of its textual content, necessary to make that content searchable and indexable in digital contexts. Yet, as I explain below, the act of producing a transcript for each video article has gradually led me to reconsider the extent to which all textuality is retroactively defined through acts of transcription.

When it came time to produce a first style guide for *JER* in 2022, I found that I needed to define certain terms, which had become more and more technical in the context of *JER*'s editorial processes: *text*, *audiovisuality*, *videographic*. What has each of these come to mean in practice? I was genuinely surprised by what emerged as our definition of textuality.

*Text:* The textual is that which can be transcribed into written language or other standardized notation. This includes the verbal content of speech and the notational content of music and dance, as well as words written in any form. Please note that the metadata of any digital object is necessarily textual, because this is (for now) how the internet works. (*Journal of Embodied Research*)

To define the textual as “that which can be transcribed” marks a significant move away from prevailing notions of language as an intrinsically distinct mode of thought. As I have argued elsewhere, drawing on black and indigenous critiques of European colonial logocentrism (Brander Rasmussen 2012; McKittrick 2021; Ferreira da Silva 2007; Spatz 2024), the assumption that language expresses a privileged mode of subjectivity — often linked to reason and rationality — is a result rather than a cause of the privileging of the written word in “western” civilisation. The act of transcribing a video article drives this point home in a practical and immediate way, as it quickly becomes obvious that “the textual” is a wildly diverse phenomenon defined only by its capacity for transcription.

While the requirements for *JER* transcripts have since been relaxed, my initial approach was completist, asking authors to include everything in a video that could be considered amenable to transcription: not only subtitles, intertitles, credits, voiceover, and recorded speech, but additionally all the bits of verbal and textual detritus that might appear onscreen, from a street sign glimpsed in the background to a name drawn in sand on the beach. Absolute transcriptive completion is impossible, yet the very attempt is suggestive for approaching the textual as a mode of thought. The more complete a transcript, the less it resembles a conventional essay (like a transcribed voiceover) and the more it becomes a kind of surrealistic poem, within which all manner of *transcribable* materials

come together in the plane of the written document. In this way, the practice of transcription makes evident just how much our ideas about language are retroactively produced by our technologies of writing. Video itself, as a medium, is not capable of making any distinction between words and gestures, speech and accent, body and place. It is only when an author or viewer attempts to transcribe words that they begin to “pop out” from the audiovisual flow, which may then find itself backgrounded through the foregrounding of text.

Situating text solely within video offers a productive deflation of the logocentric assumptions that still govern nearly all systems of scholarly knowledge. Rather than being a containing framework within which video can be located, text becomes a component of video. Taken to its extreme, this epistemic reversal can be realised in a form I call “illuminated video”: an unedited video recording that has been augmented or illuminated by textual annotations. As I said in the editorial video essay introducing *JER*’s special issue of illuminated videos: “Where an illuminated manuscript augments a primarily textual work with visual illustrations, illuminated video uses textual annotation to augment, enhance, investigate, and perhaps even critique a primarily audiovisual work” (Spatz 2021a; see also 2021b). Such an annotative approach will be familiar to practitioners of videographic film criticism, as it clearly resonates with forms like the “videographic epigraph” exercise (Keathley, Mittell, and Grant 2019). Yet in contrast to most film criticism, *JER*’s allegiance to embodied research continually foregrounds the documentary or evidentiary status of audiovisual recording, even as it departs from conventional approaches to performance documentation (Reason 2006; Sant 2017). A productive deflation of the textual as the sole legitimate mode of thought is thereby complemented by a redefinition of the audiovisual.

### On the Audiovisual

The above definition of the textual suggests an even broader complementary definition of the audiovisual as that which cannot be transcribed. According to the *JER* glossary:

*Audiovisuality*: The audiovisual refers to all kinds of audio and video content that cannot be transcribed into written

language or other standardized notation. Audiovisual elements can be *described*, as in audio description, but they cannot be directly *transcribed*. (*Journal of Embodied Research*)

The apparent openness of this definition is somewhat misleading, insofar as it is necessarily in tension with *JER*'s focus on embodied research. If *JER*'s videographic form is more than just accidentally related to its title, this can only be because the audiovisual remains bound to an operation of tracing or similarity that, even in an age of photorealistic digital image manipulation, derives its meaning from an experientially analogue relationship to bodies and places. I call this mode "experientially" analogue because it produces an image that resembles or is analogous to what the camera sees and hears; and this is the case even when the technology of recording is digital. What is the nature of that relationship today?

In recent months, *JER* has begun to receive submissions that use so-called "AI" imaging processes to produce videos in which bodies appear that never existed. While these technologies could be considered simply to extend the potential of manipulation and deception that has always been associated with photography and cinema, the shift from image manipulation to image generation puts the matter of the body under pressure to a new degree. That pressure is not linked to *JER*'s theoretical focus on embodiment, which is merely another step in a longstanding and transdisciplinary "bodily turn" of the humanities that some may consider to be already exhausted (Bradley 2023, 75–87). Rather, what is new about the use of AI image generation in the context of a scholarly video article is how it intervenes in the relationships that have historically defined audiovisuality, coming between embodiment and audiovisuality, disrupting the analogue link between body and image.

The magic, the potential, and the (often lethal) power of the camera is its capacity to "capture" or trace the world with its uniquely analogue mode of inscription. That capturing power, for all its potential violence, is what allows an infant, a tree, or a river to appear audiovisually in ways that they cannot appear textually. This mode of appearance is what I previously called the "video way of thinking." By that phrase I did not mean the work of the person holding a camera or editing a recording, but rather *audiovisual thought*: the kind of thinking that manifests itself through audiovisual appear-



ance, above all the appearance of audiovisual bodies. (On the politics of bodily appearance, see also Mirzoeff 2017.) This distinction, I maintain, only grows more ethically and politically salient with the expansion of digital image manipulation and generation. Paradigmatically, it is the difference between the thinking that occurs in front of the camera and that which occurs behind it. From the perspective of a final video “work,” these may simply be two different modes of contribution or authorship. Yet they also index radically different ways of being at stake in the recorded image. (Furthermore, while I do not have space to explore this here, the recognition of such audiovisual authorship destabilizes the assumed value of anonymity in peer review, with which onscreen appearance is incompatible.)

Is an audiovisual recording of a river really an instance of that river’s thought? Why not, if the “thought” of a philosopher is understood to be adequately traced by the precise ordering of a sequence of alphabetic letters put down in their name? Just as an aspect of a writer’s being might be captured in a written text, so an aspect of a river’s being may be captured in a video. By what right could we call one of these “thought” and not the other?

Pushing this further, the question is not only whether audiovisual recording captures the thought of those whose bodies it records, but also how that authorship might be sustained after the moment of capture. For *JER*, this is again an eminently practical matter, as the formal authorship structure of academic publishing comes into tension with the complex accrediting systems used in collaborative performing arts and filmmaking. While the latter may credit tens or hundreds of individuals in a variety of roles, authorship in the humanities rarely extends beyond two or three names. (Even in the sciences, where hundreds of names may be listed on a research paper, these usually can only be ordered rather than accorded specific roles.) Contributors and editors at *JER* face the same question again and again: Who counts as a co-author? How can we accurately acknowledge the embodied research of those whose audiovisual bodies appear onscreen? How can we escape or overturn the sedimented hierarchies according to which it is almost always the people behind the camera who are understood to be thinking, rather than those in front of it?

To clarify this point, it may help to locate *JER* near the middle of a spectrum that runs from ethnography at one end to videographic film criticism at the other. Conventionally, the visual ethnographer works across a massive power imbalance, in which they are responsible for what will become the highest status and often the only published version of the audiovisual material they collect. The ethnographer therefore carries a tremendous responsibility for that material, which may necessitate institutional ethics review processes or call for more radical approaches (see Jobson 2020). The videographic film critic, on the other hand, most often experiences the opposite power dynamic, as they stand in a relatively disempowered if potentially subversive relation to a “body” of audiovisual material that holds greater social and cultural capital. These disciplinary glosses are merely indicative (some ethnographers study powerful hegemonic cultures and some video essayists work with sensitive material), but they can serve as exemplars of the ethical and political issues that are raised when we take audiovisual appearance seriously as a mode of thought.

*JER* authors are usually working either with their own audiovisual bodies or with those of their close collaborators. These are relations of relative (or even perfect, when working with one’s own audiovisual body) equality, in which embodied audiovisual thought is carried through more or less directly into the video editing processes that produce a final work. Such relations of equality are made possible by the relative availability and affordability of video technologies. In other words, the same technological shifts that enable videographic film critics to re-edit films, and ethnographers to distribute inexpensive cameras to their participants, enable performance practitioners and other embodied researchers to produce our own audiovisual documents, making “embodied research” possible and legible in new ways. And even when *JER*’s articles do approach the relational dynamics of ethnography or videographic criticism, I would suggest that matters of ethical and political responsibility can be usefully reframed through the concept of embodied research, which foregrounds not only embodiment in general but specifically the embodiment of the researcher(s) as a central element of the methodology. It is from this perspective that *JER* approaches the space of videographic thought.



### On the Videographic

The two definitions above locate the cut between the audiovisual and the videographic in a perhaps unexpected place. If audiovisual thought refers to the onscreen appearance of bodies (whether human or not), then the videographic begins with the spatial cut between the front and the back of the camera, extending from there through the art of *videography* and into the temporally deferred practice of video editing. The double-sidedness of the camera means that audiovisual and videographic modes of thought have always been reciprocal, distinct in their practices yet dependent upon each other. Hence *JER*'s third definition:

*Videographic:* The videographic refers to the variety of media that can be incorporated within a single linear video file. In addition to audiovisually recorded material, this includes photos, drawings, sound recordings (including voiceover), diagrams, animation, and onscreen text. (*Journal of Embodied Research*)

If the textual is defined by technologies of writing, and the audiovisual is defined by the experientially analogue (even if technologically digital) technologies of recording, then the videographic today is a layered digital space in which the textual and the audiovisual come together in unprecedented ways. This is the space of montage, a space in which the "language" of cinema is continually deconstructed as new relationships are invented between diverse forms of textuality and audiovisuality. From the perspective of embodied research, there is no fundamental difference between the cinematographer, videographer, director, and video editor. All undertake videographic thought, which is always dependent upon the audiovisual as its substrate. The form and ethics of that relationship are complex and could even be understood as the primary matter explored by academic filmmaking.

For those of us with an allegiance to the critical humanities, the videographic is a space in which we face the impossible task of bringing the deep critical power of textual thought to bear in a radically different domain, one that carries some features of the textual and some of the audiovisual but cannot be reduced to either. Like a writer of words, the video editor conventionally works alone at a

desk or table, separated by their technological medium from the worlds their work is “about.” (On the phenomenology of the table, see Ahmed 2007.) The distancing effects produced by this mode of writing have on the one hand enabled the hegemonic systems of financial economics and accounting that structure capitalist and colonial modernity. But those same tools and techniques of distancing can also be leveraged for the articulation of counter-modernities, through alternative modes of writing and thinking. Historically, the cinema screen has been a very different technological space than the written page, operating on a different scale from the intimacy of reading. But as audiovisual works and written texts find themselves increasingly close companions on the screens of laptops, smartphones, and tablets (not to mention the even more interwoven textual-audiovisual assemblages of social media feeds), that difference in scale and medium begins to blur. The videographic, and the digital screen more generally, becomes a new kind of space in which juxtapositions and relations of textuality and audiovisuality might be reworked.

If the experience of *JER* has anything to contribute to an understanding of the ethics and politics of videographic thought, it is a recognition that the videographic is only meaningful in relation to the textual and the audiovisual. The moment that one can tell an app to create a video essay and have it do so is the moment of the death of the video essay — not because the resulting video could never be of interest, but because its connection to specific lineages and relationships grounded in textuality and audiovisuality will have been severed by an algorithm more opaque and impermeable than any Hollywood studio. As the videographic becomes more and more powerful, literally accruing power from the countless server farms that make up the illusory “cloud” of world computing, the urgency of sustaining textual and audiovisual relations only increases. The question of audiovisual provenance, that is the nature of the embodied and emplaced processes that generate audiovisual material, becomes more and more central to the ethics and epistemics of videographic work. These kinds of relations and matters of provenance are what I call *videographic entanglements*: the material connections that thread distant places and people together via the circulation of textual and audiovisual materials, while continually putting pressure on the meaning of those circulations. To-

day's video essayist has at their fingertips a dizzying expanse of textual and audiovisual materials that can be easily transferred into videographic space. From what perspective can and should this material be addressed, incorporated, or reworked? And how might the practice of videographic editing be different, if the materials brought into that space were treated as textual and audiovisual bodies?

With this in mind, it is incumbent upon everyone who edits video to reflect on their own positionalities and entanglements in relation to the textual and audiovisual materials with which they are working and to consider the various modes of thinking that have produced those materials. Existing legal and institutional frameworks such as copyright and intellectual property law, as well as institutional ethics review boards, are profoundly inadequate when it comes to addressing such issues. Already mired in capitalist, colonial, and patriarchal assumptions when dealing with written texts and classical production processes, such frameworks have even less to say about the rapidly intensifying circulation of textual and audiovisual bodies today. It is precisely a task for the critical humanities, in alliance with non-academic communities who also approach and appropriate filmmaking from critical or indeed politically radical perspectives, to address the ethics and politics of our growing videographic entanglements. Doing so means questioning the boundaries and relations between the textual, the audiovisual, and the videographic without collapsing them, as these are the very categories that can help us understand how we are entangled today.

## References

- Ahmed, Sara. 2007. *Queer Phenomenology: Orientations, Objects, Others*. Durham: Duke University Press. <https://doi.org/10.1515/9780822388074>
- Bradley, Rizvana. 2023. *Anteaesthetics: Black Aesthetics and the Critique of Form*. Stanford, California: Stanford University Press. <https://doi.org/10.1515/9781503637146>
- Brander Rasmussen, Birgit. 2012. *Queequeg's Coffin: Indigenous Literacies & Early American Literature*. Durham: Duke University Press. <https://doi.org/10.1215/9780822393832>
- Ferreira da Silva, Denise. 2007. *Toward a Global Idea of Race*. Borderlines 27. Minneapolis: University of Minnesota Press.

- Jobson, Ryan Cecil. 2020. "The Case for Letting Anthropology Burn: Sociocultural Anthropology in 2019." *American Anthropologist* 122 (2): 259–71. <https://doi.org/10.1111/aman.13398>
- Journal of Embodied Research. 2022. "Glossary and Style Guide." Accessed 1 May 2024. <https://jer.openlibhums.org/site/style-guide/>
- Keathley, Christian, Jason Mittell, and Catherine Grant. 2019. "The Videographic Essay: Practice and Pedagogy." Scalar. 2019. <http://videographicessay.org/>
- McKittrick, Katherine. 2021. *Dear Science and Other Stories*. Errantries. Durham: Duke University Press. <https://doi.org/10.1215/9781478012573>
- Mirzoeff, Nicholas. 2017. *The Appearance of Black Lives Matter*. Miami: Name Publications.
- Reason, Matthew. 2006. *Documentation, Disappearance and the Representation of Live Performance*. Basingstoke; New York: Palgrave Macmillan. <https://link.springer.com/book/10.1057/9780230598560>
- Sant, Toni, ed. 2017. *Documenting Performance: The Context and Processes of Digital Curation and Archiving*. London; New York: Bloomsbury Methuen Drama. <https://doi.org/10.5040/9781472588210.ch-001>
- Spatz, Ben. 2017. "Embodiment as First Affordance: Tinkering, Tuning, Tracking." *Performance Philosophy* 2 (2): 257–71. <https://doi.org/10.21476/PP.2017.2261>
- Spatz, Ben. 2018. "The Video Way of Thinking." *South African Theatre Journal* 31 (1): 146–54. <https://doi.org/10.1080/10137548.2017.1414629>
- Spatz, Ben. 2021a. "Illuminations." *Journal of Embodied Research* 4 (2). (9:20). <https://doi.org/10.16995/jer.90>
- Spatz, Ben. 2021b. "Artistic Research and the Queer Prophetic." *Text and Performance Quarterly* 46 (1-2): 81-105. <https://doi.org/10.1080/10462937.2021.1908585>
- Spatz, Ben. 2024. *Race and the Forms of Knowledge: Technique, Identity, and Place in Artistic Research*. Evanston, Illinois: Northwestern University Press.
- Spatz, Ben, Nazlihan Eda Erçin, Ilona Krawczyk, and Agnieszka Mendel. 2022. "Whiteness." *Performance Philosophy* 7 (2): 169–72. <https://doi.org/10.21476/PP.2022.72349>