Chapter nineteen

When Judgment Calls
Making Sense of Criteria for Evaluating Different Forms of Autoethnography

Andrew C. Sparkes

Evaluating research is a necessary yet contested terrain. In making judgments about any form of inquiry, we draw either implicitly or explicitly on criteria. I learned this the hard way when, some years ago, whilst walking with a colleague to the campus café for a mid-morning coffee, I mentioned to him that I was writing an article about my experiences, in my younger days, of being a high-level sports performer whose career was prematurely terminated by a serious injury and the impact this then had on my life (Sparkes, 1996). Without breaking his stride, he said, “Sounds like an academic wank to me.” His use of the vulgar British slang word for male masturbation stunned me. Not knowing how to respond, I quickly changed the subject and didn’t mention my autoethnographic aspirations to him again. But the sting of his words, flippantly released in a definitive act of judgment about work he had not seen or read, had seared itself into my body. It took me a while to find an antidote in the writings of others and to slowly cleanse the venom from my system (Sparkes, 2002, 2003, 2013a).

Whether or not we like it, as Denzin (2018) reminds us, “We can never say farewell to criteria” (p. 191), while Faulkner (2016) states, “I do not see a way out of NOT assessing the quality and effectiveness of qualitative research” (p. 665). In thinking about criteria, rather than seeing them as abstract, predetermined or universal standards to be applied regardless of context, I am drawn toward Smith and Hodkinson’s (2005) suggestion that criteria are best viewed as socially constructed lists of characteristics.

This is not a well-defined and precisely specified list; to the contrary, this list of characteristics is always open-ended, in part unarticulated, and always subject to constant interpretation and reinterpretation. . . . Our lists are invariably rooted in our standpoints and are elaborated through social interactions.

(p. 922)

Reflecting on this process, Pelias (2011) describes himself sitting at his desk contemplating what qualitative work he wants to applaud and what efforts seem lacking. He’s curious as to why he is seduced by some work but not others, why the best work seems to engage, and the weaker work seems to fall flat and leave him cold. Sitting there he is ready to consider other readings, but then he continues, putting an evaluative self forward that lists twelve contrasts between a flat piece and an engaging piece. One of these is as follows:

The flat piece, a cold dinner, is forced down, taken in with little pleasure. It lacks the heat of the chef’s passions, the chef’s sensuous self who knows, without spice, all is bland. The engaging piece makes each mouthful worthy of comment, encourages lingering, savoring, remembering. In its presence, I want to invite my colleagues and students to enjoy its flavors.

(p. 666)

The list of contrasts provided by Pelias (2011) can be seen as an articulation of the criteria he calls upon and the process he goes through when acting as an evaluative self. Others, as indicated in Table 19.1, have also generously offered lists of criteria to assist the
evocative self in passing judgment on various forms of inquiry.

More lists of criteria (as characteristics) for judging various forms of qualitative inquiry can be added to Table 19.1. To do so, however, would be to labor the point made by Sparkes (2018a) that when it comes to the criteria issue, scholars tend to create and use lists according to their specific needs and purposes. It is nonetheless so for autoethnography.

### LISTS OF CRITERIA FOR JUDGING AUTOETHNOGRAPHIES

Names have power. As Charmaz (2006) states, “names carry weight, whether light or heavy. Names provide ways of knowing—and being” (p. 396). Thus, when I act as an evocative self and review a journal article or student dissertation, I am sensitive to the names chosen by authors to describe their work as this signals their intent and purpose within the framework of a tradition of inquiry. So, when I read a phrase, such as, “in this article I draw upon an autoethnographic approach,” I begin to expect a contribution that displays the general characteristics of this genre. This is my first act of judgment.

As autoethnography has gained greater acceptance as part of its coming of age within the social sciences, various scholars have helped to clarify the key characteristics of this genre without seeking to determine what it is once and for all in an act of closure. For example, Holman Jones, Adams, and Ellis (2013) list the following characteristics: purposefully commenting on/critiquing of culture and cultural practices; making contributions to existing research; embracing vulnerability with a purpose; and creating a reciprocal relationship with audiences in order to compel a response. Likewise, Manning and Adams (2015) state that two essential qualities should be present if a project is to be considered an autoethnography.

First, any work labeled “autoethnography” should include personal experience and demonstrate, through thoughtful analysis, why this experience is meaningful and culturally significant . . . Second, this personal experience must be reflexively considered through the use of extant theory, other scholarly writings about the topic, fieldwork observations, analysis of artifacts (e.g., photographs), and/or involvement with others (e.g., interviews).

(p. 205)

More recently, Adams and Herrmann (2020), the editors of the newly established Journal of Autoethnography, in the guidance they offer to those submitting an article to this journal, state the following:

> [A]utoethnographic projects use selfhood, subjectivity, and personal experience (“auto”) to describe, interpret, and represent (“graphy”) beliefs, practices, and identities of a group or culture (“ethno”). Manuscripts published in this journal must engage these components . . . We expect every manuscript to engage at least some aspects of the “auto,” “ethno,” and “graphy,” and these components will inform how we assess manuscripts.

(pp. 2–3)
Such guidance and identifying of characteristics are useful in helping my evaluative self to judge if what is being offered to me actually is an autoethnography rather than something else, such as an autobiographical study, or a ‘confessional’ piece in which the author provides methodological reflections based on their personal experiences of being involved in the research process (Sparkes, 2020). Often, in the introduction the author tells me they have produced an autoethnography, but none of these key characteristics are present in what follows. Alternatively, the author tells me that their piece contains the characteristics named by the scholars mentioned earlier. However, when it comes to showing me these characteristics in action, none are present. My conclusion, therefore, is that whatever the piece is, it is not an autoethnography, and I am obliged to point this out in my review. This is a sharp reminder that simply claiming that one has produced an autoethnography does not mean that one has actually done so.

Of course, just how the characteristics of autoethnography outlined earlier are played out in practice is very much up for grabs (Short, Turner, & Grant, 2013). This is because there is no one definitive form or containing category of autoethnography owned by individual scholars but many variations and possibilities. Here are but a few: analytic autoethnography (Anderson, 2006), Black feminist autoethnography (Griffin, 2012), collaborative autoethnography (Toysaki, Pensoneau-Conway, Wendt, & Leathers, 2009), critical autoethnography (Boynton & Orbe, 2016), evocative autoethnography (Bochner & Ellis, 2016), indigenous autoethnography (Whitini, 2014), interpretive autoethnography (Denzin, 2014), performance autoethnography (Denzin, 2018), performative autoethnography (Spry, 2011), phenomenological autoethnography (Aguirre & Duncan, 2013), poetic autoethnography (Speedy, 2015), and psychoanalytic autoethnography (Garratt, 2014).

Not surprisingly, as Holman Jones (2005) states following her review of various definitions of autoethnography, “taking these words as a point of departure, I create my own responses to the call: Autoethnography is . . .” (p. 765). Similarly, as others have responded to this call they have, by necessity, articulated different lists of criteria for judging their preferred kind of autoethnography. For example, Hughes, Pennington, and Makris (2012) developed their notion of what makes a good autoethnography by examining those published in highly selected education journals to see how they matched the existing publication standards for empirical research established by the American Educational Research Association (AERA).

In contrast, rather than focusing on autoethnographies published in selected journals to establish which external criteria must be met, others have started with what they value and see as the purpose of the particular kind of autoethnography they aspire to before developing their criteria for judging goodness. To illustrate this point, I will now consider the lists provided by some scholars who advocate analytic, evocative, and performance kinds of autoethnography.

Analytic Autoethnography

For Anderson (2006), analytic autoethnography has the following five key features: complete member researcher status, analytic reflexivity, narrative visibility of the researcher’s self, dialogue with informants beyond the self, and a commitment to theoretical analysis. These key features, he suggests, clearly differentiate it from evocative autoethnography because the purpose of analytic autoethnography is not just about documenting personal experience, providing an insider’s perspective, or evoking emotional resonance with the reader. Rather it is about using empirical data “to gain insight into some broader set of social phenomena than those provided by the data themselves” (387). In this respect, McMahon (2016) suggests that analytic autoethnography represents a more “traditional scientific” autoethnographic approach, “with a focus both on telling readers what the tale is about and how it should, ideally, be read” (p. 307).

Drawing on Anderson’s work, others have developed his approach and the criteria used to judge both its processes and products. For example, speaking of how health researchers can produce desirable (i.e., analytic) autoethnographies that potentially contribute to advancing health-related knowledge, Chang
Like Anderson (2006), Chang (2016) also signals that her criteria, along with her suggestions for what she believes to be the most systematic process of conducting autoethnographic research, “may stand contrasted with more fluid approaches of evocative and narrative autoethnography and of interpretive and performative autoethnography” (p. 445).

**Evocative Autoethnography**

According to McMahon (2016), evocative autoethnography can also be referred to as emotional autoethnography that involves a literary approach to research that seeks to *show* rather than *tell* the reader about the subjective experiences of the author. This showing is accomplished by creating compelling stories that invite the reader to feel an emotional resonance and connection with the author, as well as gain an understanding of the culture central to the story being told.

With regard to evaluating evocative autoethnography, Bochner and Ellis (2016) offer some reflections in the form of a conversation between them and a group of students. Here, they point out that depending upon the kind of autoethnography you are doing, the criteria by which it should be evaluated will be different. Thus, if you are aspiring to do something akin to analytic autoethnography, then the criteria “should be more social scientific, such as considerations of validity, data collection, categorization processes, and generalizability across cases” (p. 212). If, however, you aspire to an evocative form of autoethnography then, they suggest, you will not be so concerned with these issues. To illustrate this point, Bochner lists the following criteria he uses for evaluating evocative autoethnography:

- I look for abundant, concrete details. I want to feel the flesh and blood emotions of people coping with life’s contingencies.
- I am attracted to structurally complex narratives that are told in a temporal framework representing the curve of time.
- I also reflect on the author’s emotional credibility, vulnerability, and honesty.
- I also prefer narratives that express a tale of two selves, one that shows a believable journey from who I was to who I am, and how a life course can be reimagined or transformed by crisis.
- I hold the author to a demanding standard of ethical self-consciousness.
- I want a story that moves me, my heart and belly as well as my head.

(Bochner & Ellis, 2016, pp. 212–213)

Likewise, Adams, Holman Jones, and Ellis (2015) created four goals for assessing the value and success of evocative forms of autoethnography. These are “making contributions to knowledge; valuing the personal and experiential; demonstrating the power, craft, and responsibilities of stories and storytelling; and taking a relationally responsible approach to research practice and representation” (p. 102).

**Performance Autoethnography**

In performance autoethnography, according to McMahon (2016), the autoethnographer/storyteller/researcher can be likened to the lead actor in their own story through a theatrical performance that “brings to life or to the stage/theatre transcribed or text-based lived experience” (p. 309). Here, she suggests, the lead actor “enacts the performance in storied form, particularly in terms of the specific cultural context within the performance” (p. 309).

Reflecting upon the possibilities of criteria for evaluating performance autoethnography, Holman Jones (2005) developed a list of actions and accomplishments she looks for in the work of others. These are as follows:

- **Participation as reciprocity**: How well does the work construct participation of authors/readers and performer/audiences as a reciprocal relationship marked by mutual responsibility and obligation?
• Partiality, reflexivity, and citationality as strategies for dialogue (and not mastery): How well does the work present a partial and self-referential tale that connects with other stories, ideas, discourses, and contexts (e.g., personal, theoretical, ideological, cultural) as a means of creating a dialogue among authors, readers, and subjects written/read?

• Dialogue as a space of debate and negotiation: How well does the work create a space for and engage in meaningful dialogue among different bodies, hearts, and minds?

• Personal narrative and storytelling as an obligation to critic: How do narrative and story enact an ethical obligation to critique subject positions, acts, and received notions of expertise and justice within and outside the work?

• Evocation and emotion as incitements to action: How well does the work create a plausible and visceral lifeworld and charged emotional atmosphere as an incitement to act within and outside the context of the work?

• Engaged embodiment as a condition for change: How does the work place/embody/interrogate/intervene in experience in ways that make political action and change possible in and outside the work?

(p. 773)

Similarly, Denzin (2018) grapples with the problem of how best to judge performance autoethnography. He offers the following criteria to help evaluate whether or not performance texts accomplish the following:

• Unsettle, criticize, and challenge taken-for-granted, repressed meanings.

• Invite moral and ethical dialogue while reflexively clarifying their own moral position.

• Engender resistance and offer utopian thoughts about how things can be made different.

• Demonstrate that they care, that they are kind.

• Show, instead of tell, while using the rule that less is more.

• Exhibit interpretive sufficiency, representational adequacy, and authentic adequacy.

• Are political, functional, collective, and committed.

(p. 197)

Of course, performance autoethnography often moves from the text to being staged in front of a live audience. Here, as Cho and Trent (2009) point out, the performance in use involves “transacting the lived experiences of Others to audiences by means of voices and bodies of the performer(s)” (p. 11). Given this change in transactional context in which, as Pelias (1992) reminds us, the performer carries the obligation to fashion an aesthetic form while the audience member has the burden of response, then modes of judgment will shift accordingly.

Regarding how to evaluate performance autoethnography when it is acted out on the stage, Cho and Trent (2009) suggest various validity criteria for guiding and developing the enactment and evaluation of performance-related qualitative work in relation to the following phases: pre-performance (criteria = imaginative), during performance (criteria = artistic), and post-performance (criteria = co-reflexive member checking). Likewise, Spry (2011), in reflecting upon what she calls performative autoethnography, has also considered the movement from the body to the page and then to the stage that involves the aesthetic process of creating performance and enacting autoethnographic analysis in front of an audience. Such movement, Spry suggests, calls for various forms of judgment associated with, for example, the self-other-context, connection, the performative-I disposition, putting flesh on the bones of discourse, practicing vulnerability as agency, and aesthetic accountability that are to be called upon at different moments.

***

The lists of criteria offered earlier for evocative and performance autoethnography do not exhaust those available for judging them. Furthermore, a careful reading of the lists will indicate that they might have some common features with the criteria associated with, for example, analytic autoethnography. This said, they also have significant differences due to the values and purposes of those producing evocative and/or performance autoethnography that need to be acknowledged and respected if judgment is to be passed on each of them in a fair and ethical manner.

THE DANGER OF LISTS

For Bochner (2000) “criteria always have a restrictive, limiting, regressive, thwarting, halting quality to them, and they can never be completely separated from the structures of power in which they are situated” (p. 269). Likewise, Adams et al. (2015) state that “evaluation criteria are political, they privilege some voices and research projects while discouraging and silencing other voices and projects” (p. 102). In relation to this, Smith and Hodkinson (2005) remind us that researchers of all persuasions, including autoethnographers, use whatever resources they have at their disposal to “support, preserve, or strengthen those...
rules (or lists of characteristics) that they approve of or are in their interests and/or to change the rules (or lists) in a direction that favors their interests” (p. 923). This is not surprising if we adopt the position held by Gingrich-Philbrook (2013) that any evaluation of autoethnography captures the efforts of real people as they deploy arguments that advance the evaluator’s own paradigm, psyche, and professional identity-work. For that matter, he adds, so do any evaluations of those evaluations, which is only to be expected since any evaluation (and evaluation of evaluations) of autoethnography always takes place in a shared world, “not at an Archimedean remove from what they evaluate” (p. 615). Thus, as Gingrich-Philbrook states, any evaluation of autoethnography is always simply another personal story “from a highly situated, privileged, empowered subject about something he or she experienced” (p. 618).

In telling another personal story, however, lists of criteria have a dark side. They can serve a strong exclusionary and legitimation function when used in a foundational, prescriptive, and normative manner to regulate the boundaries of specific forms of inquiry and control its practitioners in punitive ways. Here, lists of criteria get defined as permanent and universal to be applied to any form of inquiry under the general label of “qualitative” regardless of its intents and purposes. This list can quickly become a rigid quality appraisal “checklist” that is then used to set standards of “quality control” for all forms of qualitative research. As Gingrich-Philbrook (2013) points out, the checklist gets confused for a meta-language and something that is universally endorsed. Such lists, he argues, can too quickly get converted into a magic contract for power relations with, for example, professors marking and grading an autoethnography assignment, journal editors and reviewers making publication assessments, and administrators judging promotion and tenure applications in ways that define the absence of some criteria from the list as being a “deficit” in the work. Of course, this is not to say that checklists are worthless. Rather, the problem lies in their inappropriate application and political use especially in the hands of anonymous evaluators.

What this means is that any proposed list for judging different kinds of autoethnography can easily, as part of a magic contract for power relations, become the lists for judging certain genres at the exclusion of all others. In such circumstances the list quickly becomes ossified, inflexible, and impermeable. This creates the dangerous illusion that different genres of autoethnography are ahistorical, fixed, and bounded with predetermined and separate judgment criteria for each one that have no connection with each other. That is, one either does analytic autoethnography or one does evocative or performance autoethnography in accordance with specific lists of criteria for each genre. Such a view is highly problematic.

As Winkler (2018) points out, the first two criteria proposed by Anderson (2006) for analytic autoethnography could also be used to describe evocative autoethnography, a point acknowledged by Anderson himself who does not state that creative or evocative writing must be excluded from analytic works. Likewise, in seeking to expand analytic autoethnography and enhance its potential, Vryan (2006) notes that including data from and about others (Anderson’s third criteria) is not a necessary requirement for all analytic autoethnography, and that the necessity, value, and feasibility of such data will vary according to the specifics of a given project and the goals of its creators. Furthermore, Vryan argues, a distinction between analysis and creative or evocative first-person writing styles is “unnecessary and counterproductive, as are implications that an analytical project must avoid delving too much, or too expressively or exclusively in the autoethnographer’s experience” (p. 407). Significantly, in reflecting back on the position he took in 2006, Anderson states that he has since developed a more nuanced and appreciative understanding of a range of potential analytic autoethnographic styles, and that if he were to remain committed to an analytic model then he would do with a “greater sense of blurred boundaries as opposed to clear distinctions” (Anderson & Glass-Coffin, 2013, p. 64).

Analytic autoethnography, therefore, can include evocation and emotionally rich texts and this is not incompatible with analysis. Accordingly, Hayler (2013), Sparkes (2020), and Winkler (2018) reject the notion that evocative and analytic autoethnography need be mutually exclusive and note that a fruitful combination can enlarge the spectrum of autoethnographic research and open up further possibilities for development. Regarding these possibilities, Colyar (2013) shows how, in combining expressive writing (that foregrounds energy and emotion) with transactional writing (that emphasizes analysis and theorizing), along with poetic writing (that creates literary texts), she is able to produce an evocative analytic text. All this leads Tedlock (2013) to call for the braiding of evocative with analytic autoethnography in ways that produce “powerful writing about the self in the world in order to help change the world” (p. 361).

Tedlock’s (2013) notion of braiding becomes even more important when different forms of autoethnography are taken to exist on a continuum rather than in discreet boxes. For Tullis (2013) autoethnography “exists on a continuum from highly fluid and artistic...
to formulaic and highly analytic” (p. 245). Likewise, Manning and Adams (2015) identify the four common orientations used by many autoethnographers to design, conduct, represent, and evaluate their projects. These are as follows: social scientific autoethnography (i.e., analytic autoethnography in Anderson’s [2006] terms); interpretive-humanistic autoethnography; critical autoethnography; and creative-artistic autoethnographies. Importantly, Manning and Adams point out that although they list four distinct orientations, “it is not unusual for autoethnographers to blend the goals and techniques of each in a single project or as they write about the same experiences over time” (p. 191).

For Ellis, Adams, and Bohner (2011), therefore, autoethnography, as a method, is necessarily disruptive of binary oppositions and can be rigorous, theoretical, and analytical as well as emotional, therapeutic, and inclusive of personal and social phenomena. All this suggests that the boundaries of autoethnographic types or orientations in terms of the balance between showing and telling, and the lists of criteria associated with them, need not be viewed as rigidly bounded systems but as something more permeable and fluid in nature, which allows them to be used in a creative rather than a conservative and formulaic manner.

Significantly, a number of the scholars cited earlier are acutely aware of the dangers associated with their lists of criteria being used inappropriately for judging autoethnographic inquiry. For example, Tracy (2010), reflecting on the possible uses of her own list warns that “grasping too strongly to any list of rules—and treating them as commandments rather than human made ideas is an act of delusion, suffering, and pain” (p. 849). Likewise, Adams et al. (2015), Hughes et al. (2012), and Schroeder (2017) emphasize that their lists should not be taken as a universal, closed, or permanent set of criteria against which to measure all autoethnographic texts regardless of purpose or context. They offer these criteria as starting points for thinking about how to appraise the works of arts-based research. While their criteria may act as a common point of reflection, Barone and Eisner do not want them to be seen as a fixed recipe that all must follow as this would lead to rigid standardization at the cost of innovation. They invite researchers, therefore, to use their own judgment in applying their proposed criteria to examples of arts-based research and Barone and Eisner urge researchers to use their imagination in ascertaining other criteria that may emerge from encounters with arts-based work in the future.

In generously providing their own lists of criteria for the consideration of others as starting points and cues for perception, the aforementioned authors are willing to describe what one might do but are not prepared to mandate what one must do across all contexts and on all occasions prior to any research being conducted. In this sense, therefore, we can begin to discuss the characteristics of a particular approach to inquiry, such as analytic, evocative, or performance autoethnography and simply note that these criteria are the way different researchers seem to be conducting and evaluating this particular kind of autoethnography at the moment and that this could, and probably will, change over time.

**CREATIVELY CREATING LISTS OF CRITERIA**

Smith and Deemer (2000) emphasize that any list we bring to judgment is always open-ended, and we have the capacity to add to or subtract characteristics from the lists. This is necessarily so because the criteria used to judge a piece of research can change depending upon the social, cultural, political, and historical contexts in which they are called upon to differentiate the “good” from the “bad.” There is always the possibility, especially when something novel comes along, that one must reformulate one’s lists and replace the exemplars one calls upon in the “never-ending process of making judgments” (p. 889). Regarding this possibility, Smith and Hodkinson (2005) note that the limits for recasting lists derive less from theoretical labor and more from the ways in which they are worked and reworked within the context of actual practices or applications.

That the creation and reworking of lists of criteria are accomplished in the doing and engagement with actual inquiries rather than via the distillation of some abstracted epistemology is evident in the work of Holman Jones (2005). Speaking of the list of criteria she has developed for judging her own work and that of others, she notes that they are changing and

---

I wish to articulate what I like and what I don’t without imposing my evaluative stance but acknowledging that I have one that guides my practice as a reviewer, teacher, and writer. I leave open the possibility of other evaluative and more productive schemes.

(p. 666)

Barone and Eisner (2012) also emphasize that each of the criteria they have included in their list for judging arts-based research function as cues for perception. They offer these criteria as starting points for...
“are generated in the doing of this writing rather than outside or prior to it” (p. 773). Likewise, Gingrich-Philbrook (2013) argues that lists “make so much more sense as something developed over time and experience, something that changes and grows, adapts to different writers, writing different projects, for different purposes, at different times” (p. 619). Indeed, even when discussing alleged universal criteria for judging qualitative research, Tracy (2010) acknowledges that understanding qualitative goodness is best appreciated by embodying the methods ourselves as apprentices in the practice of research and also vicariously studying the dilemmas of others.

An example of a researcher engaging with lists proposed by others, prior to creating a bespoke list of criteria to guide her own work is provided by Le Roux (2017). As part of her deliberations on the notion of rigor in autoethnographic research, Le Roux conducted a literature review of relevant research in which established autoethnographers gave their views on the nature of this genre and the issue of rigor. Le Roux also called upon her own experiences of undertaking an autoethnographic study, and finally, she drew upon data generated from a questionnaire administered to proven researchers in her own institution to access their perceptions of the rigor of autoethnography as a research method. In this process, she recognized that the criteria she had applied to her own autoethnographic study were more inclined toward theory-driven, analytical research and that she had not evaluated her research in terms of it being a reflexive, honest account of her own experiences situated in culture.

As a consequence, Le Roux (2017) realized that she had not used criteria such as resonance, researcher subjectivity, narrative truth, reflexivity, aesthetic merit, or plausibility as advocated by, for example, Bochner and Ellis (2016), Manning and Adams (2015), and Richardson (2000). Le Roux then goes on to develop the following list of five criteria, each of which has inbuilt to them the expectation that the research is ethical.

- **Subjectivity**: The self is primarily visible in the research. The researcher reenacts or retells a noteworthy or critical personal relational or institutional experience—generally in search of self-understanding. The researcher is self-consciously involved in the construction of the narrative which constitutes the research.
- **Self-reflexivity**: There is evidence of the researcher’s intense awareness of his or her role in and relationship to the research which is situated within a historical and cultural context. Reflexivity points to self-awareness, self-exposure, and self-conscious introspection.
- **Resonance**: Resonance requires that the audience is able to enter into, engage with, experience or connect with the writer’s story on an intellectual and emotional level. There is a sense of commonality between the researcher and the audience, an intertwining of lives.
- **Credibility**: There should be evidence of verisimilitude, plausibility, and trustworthiness in the research. The research process and reporting should be permeated by honesty.
- **Contribution**: The study should extend knowledge, generate ongoing research, liberate, empower, improve practice, or make a contribution to social change. Autoethnography teaches, informs, and inspires.

Having constructed her own list, Le Roux (2017) makes the point that such checklists cannot substitute for informed judgment and that any appraisal of an autoethnography should be subject to individual judgment based on insight and experience. For her, competent researchers and appraisers of research “must acquire not only the ability to use and understand the application of various research skills but also the acumen to judge when some kinds of research are likely to prove more productive and germane than others” (p. 204). This raises the pedagogical issue of how we might go about assisting students, colleagues, and ourselves to develop the ability to make informed judgments about different kinds of autoethnography.

### Lists of Criteria and Their Pedagogical Potential

Newcomers to qualitative inquiry can be bewildered by the vast array of criteria available for judging their work. For Tracy (2010) such bewilderment can be reduced initially by offering students her eight universal criteria as this provides them with what she calls a “common language of excellence for qualitative research” (p. 849). Equally, the lists of criteria provided earlier for judging different kinds of autoethnography can also reduce bewilderment by providing a common language or set of characteristics for discussing what goodness might mean in each kind.

As Le Roux (2017) states, given that autoethnography can be approached from diverse orientations, having available a concise list of criteria can be useful. The usefulness of lists is evident in a collaborative project described by Schroeder (2017) that drew...
together a group of scholars in the field of Library and Information Science to explore, via autoethnography, an approach they were all unfamiliar with, the issue of the self as subject in the kinds of research they conducted (see Deitering, Schroeder, & Stoddart, 2017). Schroeder admits that at the start of this project he had no idea what autoethnography was and so he undertook a major literature review to explore what criteria were possible for reviewing or evaluating autoethnographies. Based on this review Schroeder pulled together a number of criteria around the following general categories: Revealing the Self (auto); Exploring Culture/Society (ethno); Storycraft (graphy); Ethics; Social Justice and Transformation. He notes that some of the criteria fitted into multiple categories but that he just slipped them in where it seemed right. This was because categorization was not a goal in itself but rather a way to talk about the disparate criteria that he found.

Adopting Gergen’s (2014) view that local communities of qualitative scholars are best positioned to create their own criteria to help them review and evaluate their work so as to create better research, each chapter in the edited volume was reviewed by another contributing author and one of the editors. To assist this process, Schroeder (2017) distributed his list of criteria to all involved.

We used the list of criteria I gleaned from my readings as a starting point. Each author picked criteria from the list, ones that resonated with the goals they had for their own chapter. They were encouraged to change any of the criteria and to invent new ones as needed. The list they individually created was the criteria that the reviewers used to help make sure they met their goals.

(p. 324)

Once the reviews of the chapters had taken place, Schroeder (2017) then surveyed each of the author-reviewers with various questions about whether or not they found his list of criteria helpful in developing and improving their own autoethnography prior to having it judged by others. They were also asked if the criteria selected by authors to evaluate their work from the list provided help or hindered their review. All of the author-reviewers said that having the criteria to choose from helped them focus more precisely on their task, and that, after using their criteria they also felt good. One stated, “It was comforting to be able to communicate directly to my reviewers what I hope to achieve with my writing.” Another commented, “The criteria provided comfort that I had in fact actually written an AE.”

Schroeder (2017) notes that none of the authors modified any of his criteria, but they did add some questions or concerns of their own to the list. One felt that “using criteria, instead of questions, likely encourages a more robust and critical response from a reviewer/evaluator/reader.” Another pointed out how the criteria, especially with respect to autoethnography, made the reviews feel less of a critique:

The process was less about evaluation, in the end, and more about creating a conversation about perceptions of the draft. I think this is particularly useful for AE writing, where at times the subject matter might be rather personal, and a reader/reviewer may hesitate to critique or question the subject matter of the author’s approach. The evaluative criteria create a sort of formal layer of mediation—it gives both the reader and the writer a comfortable space where critique can happen without concerns related to sensitivity about the subject of the AE.

(Schroeder, 2017, p. 325)

Importantly, the responses to Schroeder (2017) also indicated that those involved found the use of criteria to be equally useful in reviewing another author’s work. As one stated, “I was able to focus not just on a review of the overall piece but also on what the author herself indicated she hoped to accomplish in her work.” This supports the previous comment that talked about how the criteria helped make the peer review more of a supportive conversation than a traditional critique. Thus, a virtue of using criteria might be that it can move us beyond evaluation toward a way of reseeing the text.

As the feedback provided to Schroeder (2017) suggests, lists of criteria, when used as starting points, can provide an initial sense of security and direction for researchers when they take the risk and engage with autoethnography for the first time. In this instance, Gingrich-Philbrook (2013) notes the following:

Budding autoethnographers may very well want the reassurance of a checklist outlining things a good autoethnography does, the qualities it possesses, because that might help them decide when they have finished a piece they’re working on. Wouldn’t it be great to have a kind of a cross between an existential oven-timer and a drag-queen fairy godmother to look over your shoulder at the screen and say “Bing! You’re done, Honey: this shit is baked; anyone tells you different, I will come over and stomp their ass”?
Of course, it is not only budding autoethnographers who need such reassurance. I suspect that many a seasoned scholar has wished for, and found in some guise or other, the existential oven-timer and drag-queen fairy godmother described by Gingrich-Philbrook (2013). I certainly know I have and still do. At times, we all need somebody we trust and respect to say, “You’re done, Honey; this shit is baked.” And equally, sometimes, we need this very same person to gently tell us that, “You’re not quite done yet, Honey; this shit is still half-baked and not ready for public consumption.” Indeed, as a teacher developing the confidence of my students on qualitative courses and in supporting my colleagues when they engage with creative analytical practices, I have often adopted, sometimes knowingly and sometimes less so, the role of existential oven-timer and drag-queen fairy godmother. It is a worthy role to be celebrated.

Even when not present in corporeal form, the combined existential oven-timer and drag-queen fairy godmother can manifest itself in a list of criteria. For example, Gordon and Patterson (2013) acknowledge how Tracy’s (2010) list provided them with a useful guide for analyzing and evaluating their own work framed by womanist caring theory and suggest that her criteria could also prove useful with other theoretical frames depending on the intentions and purposes of the studies involved. They propose that when writing qualitative studies for publication, Tracy’s criteria can provide a tool for scholars to monitor the quality of their own work and they believe that scholars “will strengthen their work if they make their use of Tracy’s criteria explicit” (p. 693). Of course, any of the lists provided earlier can prove equally useful guides for the tasks described by Gordon and Patterson for other researchers depending on their starting points, intentions, and purposes.

But then, I begin to worry a little about the notion of criteria as a tool, with its mechanistic, linear, and functional implications, to strengthen autoethnography as a process rather than a product. My undergraduate and postgraduate students often ask me what criteria I use as reference points when I go about writing an autoethnography. They feel unsettled when my answer is “None.” A sense of unease is also evident when I tell them that I have never produced an autoethnography with a view to it being of a certain kind, be it, analytic, evocative, performance, or any other.

By way of explanation, I offer them the words of Winterson (2012) who draws attention to two kinds of writing: “the one you write and the one that writes you. The one that writes you is dangerous. You go where you don’t want to go. You look where you don’t want to look” (p. 54). I also ask my students to access the sublime words of Leonard Cohen in his 2011 How I Got My Song Address at the Prince Asturias Awards in Spain. Here, Cohen tells the audience that he feels uneasy because he has always felt some ambiguity about an award for poetry. This is because, for him,

Poetry comes from a place that no one commands, that no one conquers. So I feel somewhat like a charlatan to accept an award for an activity that I do not command. In other words, if I knew where the good songs came from I’d go there more often.

Echoing such thoughts, I inform my students about my feeling that my own autoethnographic stories have always written me far more than I have ever written them as part of an embodied process rather than just a textual product. Thus, as I have suggested elsewhere (Sparkes, 2013b), autoethnography is at the will of the body, often involving unbeknown yet-to-be-told stories that circulate within us at the pre-objective, fleshy, multisensory, and carnal level, not yet ready for language to take its hold. When the body is ready to release its story, it lets us know in subtle ways so that we can accept its gift and engage in the sensuous somatic work of crafting a tale for the telling to self and others (Sparkes, 2017).

Of course, this then leads to the question of how, acting as an evaluative self, I pass judgment on autoethnographies produced by others. In response, I tell the students that for me this is not a purely cognitive, linear, or rational act but rather a messy, tentative, contingent, sensuous, and deeply embodied process in which my evaluative self feels its way into the autoethnography in front of me, reading it multiple times with gaps in between where a cocktail of thoughts and emotions mingle in my body as I drift toward certain kinds of judgment calls over others. Then, I try and show the students this messy process in action by sharing some published autoethnographies with them and offering my reflections on the criteria I am drawn toward in passing judgment on each as an evaluative self (see Sparkes, 2020). In doing so, it becomes evident that I draw upon multiple criteria from the various lists available for judging different kinds of autoethnography as well as criteria beyond these lists. Hopefully, this illustrates to the students that while lists of criteria are useful as starting points, they are not enough on their own and so the evaluative self must be creative in using them along with any other criteria that are relevant to making a fair, balanced, and ethical judgment about the quality of a piece of work.
Given what I had said earlier, it is important for me that when it comes to judging the products of autoethnography my students are invited to think about and with the various lists of criteria available that are often contested, overlapping, and contradictory. I ask them to reflect on how they feel about any given criteria in their guts and in their flesh. They can then start to consider the ways in which this informs how they make what Beckett and Hager (2002) call “embodied judgments” that are practical, emotional, and corporeal as well as discursive in nature. As Svendby (2019) articulates in her PhD thesis:

There are so many lists, so many different opinions, and views on criteria. . . . I am drowning here! Which one am I to choose? How can I even begin to legitimise my choice of one in particular and not another? Should I put together a new list based on other lists? No. The “not’s” are yelling that it is not a good idea. Hmm. . . . I think I will have to rely on my emotions in this case. Yes. I feel a “do” coming on. I will use the list that I am most drawn to, the one that feels . . . if not right exactly, then at least less wrong. . . . Yes. That feels right. I feel the “do.”

(p. 137)

In terms of “feeling the do.” I therefore invite students in my classes to construct their own list of criteria from existing lists as part of a braiding process and to create and add their own criteria if they so wish. In this process, they can begin to explore why they are drawn toward, how they feel about, and are seduced by some criteria rather than others. I then ask them to consider if it would be fair and ethical to apply their preferred criteria to all forms of autoethnography regardless of the different intents and purposes others might have for their work. We also reflect on what might happen if an autoethnography that self-defined as evocative was sent for review to a person whose preference was for analytic autoethnography with its associated criteria, and vice versa. All of which leads to a consideration of what Gadamer (1995) calls the effective histories and the prejudices each person brings to the selection of criteria and how they are used in judging their own work and that of others.

As Smith and Deemer (2000) remind us, in any encounter with a production, especially something different or “new,” one must be willing to risk one’s prejudices. They point out that, “Just as in the process of judgment one asks questions of the text or person, the person or a text must be allowed to ask questions in return” (p. 889). Approaching something novel or unfamiliar, therefore, requires that one be willing to allow the text to challenge one’s prejudices and possibly change the criteria one is using to judge the piece, thereby changing one’s idea of what is and is not good inquiry. This said, Smith and Deemer point out that to be open does not mean to accept automatically, and that one may still offer reasons for rejecting something new. The outcome of any judgment is uncertain. They also stress that there is no method for engaging in the risking of one’s prejudices. If anything, Smith and Deemer argue that “to risk one’s prejudices is a matter of disposition—or, better said, moral obligation—that requires one to accept that if one wishes to persuade others, one must be equally open to be persuaded” (p. 889). This view is supported by Gingrich-Philbrook (2013) when he states: “To evaluate autoethnography in a genuinely useful way, you have to open yourself up to being changed by it, to heeding its call to surrender your entitlement” (p. 618).

Risking one’s prejudices and surrendering one’s entitlement within the magic contract for power relations in relation to judgment criteria for autoethnography, or any other form of inquiry, is no easy task. It means assuming the responsibility to listen carefully and respectfully, attempting to grasp emotionally, viscerally, and discursively what is being expressed in something “different” so that judgment might be passed in an ethical, fair, and caring manner. This requires the qualities of connoisseurship as described by Eisner (1991). For him, connoisseurship involves the ability to make fine-grained discriminations among complex and subtle qualities, it is the art of appreciation and can be displayed in “any realm in which the character, import, or value of objects, situations, and performances is distributed and variable” (p. 63). Eisner emphasizes that the term appreciation should not be conflated with “a liking for” since there is no necessary relationship between appreciating something and liking it. For him, “nothing in connoisseurship as a form of appreciation requires that our judgments be positive. What is required (or desired) is that our experience be subtle, complex, and informed” (pp. 68–69).

In seeking to develop the characteristics of connoisseurship in myself and my students, I want to make it clear that this does not involve what Smith and Deemer (2000) call a romanticized “intellectual flight from power” (p. 202). Part of connoisseurship requires a critical awareness and appreciation of how power and politics at various levels operate and are interwoven into the complex social interactions that define which criteria, from all those available, are selected to sort out the good from the bad at a given historical moment.
Lists of criteria, as pedagogical devices, can assist students to explore issues of power and politics in relation to how they are created, legitimized, and used to foreground certain voices and silence others. To this end, I share with students my own experiences of crafting an autoethnographically informed piece of work that spoke truth to power and the consequences that followed when, as a hostile reaction to this work, managerial power was enacted in its most raw, intimidating, and questionable form that led to me leaving a university I had worked in for twenty-two years (see Sparkes, 2007, 2018b). Against this backdrop and given that any list of criteria is never neutral in its construction or its use, I encourage and help students to develop their skills in the darker arts of conceptual self-defense and strategies of self-preservation. For example, calling upon various lists of criteria, we consider how to mount an articulate response to my former colleague’s charge that autoethnography is an “academic wank,” or as others might state it more politely, masturbatory or self-indulgent.

Using lists as a way of learning to play the criteria game for me, therefore, is not an act of consent to dominant views of what constitutes good or bad research. Rather, as Tracy (2010) notes, it is a strategically designed way to respond and act within, rather than being “worked over” in hostile situations. All this said, I am fully aware that questions about how, as qualitative researchers of any kind, we create and construct our lists of criteria and the uses we put them to in various contexts will not be found in epistemology. They will, however, as Smith and Hodkinson (2005) remind us, “be found in our reasoning as finite practical and moral beings” (p. 950). And so, the conversations will necessarily continue.

REFERENCES


Hayler, M. (2013). When we got to the top of Elm Road. In N. Short, L. Turner, & A. Grant (Eds.), *Contemporary British autoethnography* (pp. 17–32). Sense Publishers.


Svendby, R. (2019). *Becoming the ‘other’: A qualitative study of power, masculinities and disabilities in the lives of young drivers after road traffic accidents* [PhD thesis], Faculty of Medicine, Oslo University.
