Abstract

Since the replication crisis began in 2011 there have been a number of heated controversies, usually over the results of replication studies. In these controversies, accusations of bullying, bad faith, a lack of civility and professional courtesy, and other moral flaws are regularly exchanged. A ‘tone debate’ has developed, recognised as a specific, ongoing discussion in and with the reform movement in psychology. The tone debate is not merely about tone and manners: in debating how to debate, psychologists explore the connections between knowledge, power, and subjectivity. In what is at first sight a marginal ‘meta-debate’, and is in fact considered as such by some of the psychologists involved, fundamental issues are at stake. What is the place of criticism in science, and how to do criticism? How to be both critical and collegial? Is it possible to separate the science and the scientist? I will show that the tone debate has brought an increasing focus on the virtues of the scientific self. It is no longer enough to follow procedures: being good is as important as doing the right thing. I will argue that this transforms relations between knowledge, power, and subjectivity in science.
1. The topic of my talk is the ‘tone debate’, which is a debate within and with the reform movement/community in psychology. (‘Tone debate’ is a term the reformers use themselves; it is ‘a thing’ for them.)

The tone debate is a meta-debate, because it is a debate about how to debate: how to criticise without hurting people, how to be critical and civil at the same time, should we be civil at all, should ‘criticism’ in fact be central to scholarly communication etcetera. It is a debate primarily about criticism, but it is part of a wider discussion about a more general question: how should we relate to each other in academia, in science? How should we get along? For example, how should we relate to students, to mentors or mentees, should peer review remain anonymous or not, is ‘this is more of a comment than a question’ acceptable after a presentation etcetera.

I will point out two themes in the tone debate.

First, power is discussed a lot in the tone debate: power differentials in science, the rights and responsibilities of early career versus established researchers, punching up versus punching down. Who is in control, who controls the debate, should the debate be controlled?

Secondly, the tone debate has brought increasing attention to ‘the scientific self’, in addition to the behavior of scientists; attention to what or how a scientist should be, as well as to what she should do; this theme one could call subjectivity, or ethos.

And thirdly, although this is less obvious, the tone debate is also a debate about epistemology; how to criticise, and more generally how we get along with each other has repercussions for the knowledge the discipline produces.

I will argue that the tone debate offers the possibility of redefining the relations between power, subjectivity, and truth. That is why, in my opinion, the tone debate is very important.

2. First I want to point out that the tone debate has precursors. There have been discussions before about how to discuss, how to criticise, how to get along as scientists, philosophers, intellectuals, as people whose primary interest is truth. I will very briefly describe two practices and the debates about them, which can serve as backdrop, as two models to compare the current tone debate with. [slide]
The first is the 18th century culture of the salon and its ideal of polite conversation. A salon was tightly regulated, governed by the hostess, the salonnière. The constraints were considered necessary because 18th century French intellectual life was very competitive, it was “both militant and personal” (Goodman, 1994, p. 91). “The problem facing the philosophes was deceptively simple: how were they to continue to debate and disagree in person and still remain collaborators and friends?” (Goodman, 1994, p. 96) This problem became acute when they increasingly found themselves together in Paris, debating in person rather than by letters. To manage the strong egos with their clashing views, the salonnières watched over the conversations, and if you spoke too long or you were rude the salonnière would cut you off. (What we would now call ‘tone-policing’.) The politeness they imposed created equality and protected the weaker voices against the louder ones. But the ideal of polite conversation was contested. Rousseau for example complained about the censureship of the salonnières and wanted a more ‘masculine’ conversation. Others insisted, however, that “disputes have only ever produced confusion, obstinacy, and hatred” (Prémontval cited in Shea, 2010).

An example from antiquity is the concept of parrhesia, a Greek word meaning literally ‘to say everything’ – so: to speak freely, frankly, bluntly, boldly. It was unadorned, plain speaking, the opposite of flattery. Originally parrhesia referred the right of every free citizen in Athens to speak their mind in the Assembly (the ekklesia), but it remained a cherished value in other contexts right until the early Christian period. That is a period of about eight centuries in which parrhesia was recognized as being of great importance, but it was also potential danger. It could be abused, or perhaps better: what looked like frank speech was not always truly parrhesia. For example, the Cynics (Diogenes, Crates and so on) were seen as paragons of parrhesia. They were admired as philosophers who had the courage to speak truth to power (Diogenes telling Alexander the Great to get out of his sun), but they were also criticized as just being rude, uncivilized, and obscene.

Two themes were central to both debates: power and subjectivity. The power of the salonnière, for example, and her specific qualities as a woman. The lack of power of the Cynic, and his ruthless honesty. In the current tone debate power and subjectivity are equally central.

---

1 Michel Foucault studied the history of parrhesia in the last few years of his life, and my take on the tone debate is inspired by his work. See in particular (Foucault, 2009).
the Association for Psychological Science, written by prominent social psychologist Susan Fiske. There is a lot to say about this text, its context, and its history, but I will focus on those themes of power and subjectivity.

Firstly, Fiske objects to the ad hominem character of the criticism. The critics attack the person as well as the content, and this is wrong. They must be kept separate. “Personal insults are not scientific discourse.” (Fiske, 2016) This is about the role of subjectivity in academic discourse, and Fiske basically says: it has no role. At the same time, it is a power issue. Fiske accuses the critics of violence, of harassment. To her, the two are connected: it is bullying because people are named, in public. Her own text on the other hand did not constitute harassment because she didn’t name names (Letzter, 2016).

Secondly, Fiske, one could say, objects to the parrhesia of the critics as such – to the fact that they take the liberty to speak freely and frankly in public. This is also wrong. Scientific debate should not be free but moderated, monitored, curated, reviewed. This, again, is about power, about the need to control public scientific debate.

So Fiske proposes an ideal of academic debate not unlike the polite conversation of the salon. Everyone politely discusses the matter at hand without upsetting anyone, and no one raises their voice. The hosts of the conversation (reviewers, editors, moderators) monitor the proceedings and intervene when things get out of hand.

The response to this and other tone criticism has been diverse. I think a number of arguments, positions, proposals can be distinguished, but they do not form a package. The reform community is diverse, and tone is a matter of debate within the community as well.

One response has been to say that there is no tone problem, that there may be isolated instances of bad behavior, but not a pervasive culture of bullying (Chambers, 2016; Yarkoni, 2016). Secondly it has been argued that the tone debate is a distraction that shifts our attention from the real problem (the crisis in psychology) to a side issue. It is seen as a tactic of the powerful in the discipline, people such as Fiske, eager to protect their legacy and their status (Chambers, 2017; Srivastava, 2018a). [slide] “Always ask who has the power. When a rich, tenured, popular figure at a famous university is claiming to be bullied, always ask first if what they mean instead is closer to ‘I feel disrespected’ or ‘get those kids off my damn lawn’.” (Heathers, 2018) “(T)he tone discussion is being used
strategically by purveyors of bad science to maintain their power” (Gelman, 2018). “Tone policing” is a way of guarding the status quo (Brown, 2019; Roberts, 2019).

A third response has been to acknowledge that tone is important, or at least that good manners, or etiquette, are important in debate. These good manners are then typically specified in lists of principles, or steps to be followed. A commonplace in these lists is that in scientific debate one should not make the matter personal: use “language that focuses on the ideas rather than the authors” (LeBel, 2014), “criticise the science, not the scientist” (Bishop, commenting on Gelman, 2016). Here for example are Uri Simonsohn’s (2016) three ideas for “civil criticism” [on slide]. This is the same separation that Fiske insisted on.

However, this separation between science and scientist, this strict prohibition of the ad hominem, cannot always be maintained. We should not immediately assume malice, but if there is “water-tight evidence” that people are intentionally engaging in questionable research practices they should be called out (Bishop, commenting on Gelman (2016)). “There are times when we actually want to talk about the researcher” (Scheel in Sho, 2016). So the ‘no ad hominem’ principle should be nuanced: “We should have a low bar for talking about science and a high bar for talking about scientists” (Srivastava, 2018b). (See also Dahly (2018) and Dahlgren (2018) about ad hominem.)

A fourth type of response to the tone issue is to formulate virtues of scientists: how scientists should be, rather than how they should act; ethos, rather than manners. Often mentioned virtues are civility and humility, intellectual humility specifically. Some examples: “Try not to centre yourself as the person who is right, but as someone with intellectual humility who has some questions.” (Simpson, 2019) And: “part of intellectual humility is cultivating a sense of interest (at least) for results counter to expectation/desire.” (Nosek, 2019) So humility must be cultivated, and in fact the crisis itself could be said to have fostered humility: “the replication crisis highlights the operation of psychological science at its best, as it reflects our growing humility.” (Lilienfeld, 2017, p. 660)

In other words, a kind of self-transformation is required. We find accounts of researchers who have gone through such a transformation, who had a kind of epiphany and changed their attitude and their practices because of it. An example: a researcher writes a blog post describing what he ironically calls his “Methodological Awakening (tm)” (Gervais, 2017). It was brought on when another team of researchers replicated one of his experiments, a “small, cute, barely significant experiment”, and couldn’t reproduce the effect. It was a life-changing experience. One commenter called the post “amazing”, another wrote it
was “downright inspirational” (Calin-Jageman and Funder, in Gervais, 2017). Similar conversions have been described by Uri Simonsohn (2016), and by Michael Inzlicht (Inzlicht & Inbar, n.d.), who admonishes his colleagues that “the only way we can really change is if we reckon with our past, coming clean that we erred; and erred badly.” (Inzlicht, 2016)

There are initiatives to facilitate such confessions, such reckoning with the past: the Loss of Confidence project, for example, intended to “create an environment that incentivizes self-correction” (Rohrer et al., 2018, p. 2). Researchers were invited to submit short statements in which they described a finding they no longer believed in, i.e. a result they had lost confidence in. The project drew a lot of attention and praise, but only six statements were submitted. A similar celebration of humility is Veronika Cheplygina’s ‘How I fail’ series of interviews with researchers, which she publishes on her blog. To Cheplygina, ‘failures’ include rejections of papers and grant applications, but also personal, moral failings like bad mentoring (Cheplygina, 2019). Most interviewees restrict their answers to the first, more easy to acknowledge kind. But these failures too can have life changing, self-transforming impact. As one interviewee said of a particularly negative set of reviews: “I look back on that grant now and credit the reviews with shaping my perspective on granting and writing as well as showing me that good scientists must be humble and persistent. Humility and persistence were my two biggest lessons.” (Mike Yassa, in Cheplygina, 2017)

Finally, tone is increasingly discussed in terms of equality, inclusiveness and diversity. For instance, the term ‘bropenscience’ has been used to refer to what is seen as an aggressively masculine culture of criticism in the name of science and rationality. Power is again the theme here. “Competitive discourse tends to hurt women and help men, in terms of how they are perceived by others.” (Ledgerwood, Haines, & Ratliff, 2015) We need ‘critique’ instead of ‘criticism’ (Devezer, 2019; Rooij, 2019). We should stop “trashing others” in the name of rigor (Heemstra, 2019). Perhaps we should even stop speaking of a crisis, which only invites “panic and overreaction”, and instead aim at “thoughtful and careful improvement” (Navarro, 2019). ‘Kindness’ is often put forward as the alternative (e.g. Lapan, 2019). Of course, the use of the term ‘bropenscience’ itself has been critcised for being a slur that is used to silence people (Jussim, 2019).

To summarise: in response to criticism about tone, psychologists in and outside the reform movement have reflected on power and on the scientific self. Both in the interest of truth: to reform science, it is not enough to pre-register, enlarge our samples, and become bayesians. We must fight institutional power (the

---

2 Daniël Lakens, however, shared the story of his dubious first publication.
3 Introduced by @o_guest
dominance of institutionally powerful members of the discipline, academic hierarchy in general, the power of the publishers), and we must improve ourselves and the way we deal with each other, ridding our relations of dominance. That is not to say that everyone agrees about power and subjectivity, but these two issues are central in the debate. Even those who reject the tone debate altogether do so with arguments about power and the scientific self: for example, that the tone debate plays in the hands of the powers that be in the discipline, that we shouldn’t be thin-skinned and learn to accept criticism as well as praise (Chambers, 2017). (This too is a call for self-transformation.)

5.

Writing this paper I have struggled with the question what my point should be, my message, because the debate itself is already rich in messages, calls, proposals, and statements. But perhaps my historical/philosophical perspective offers a few suggestions.

Firstly, I think the tone debate is important, because it is a debate that forces us to reconsider the relations between power, subjectivity, and truth in Psychology. In the issue of tone epistemic issues connect with ethical and political questions. The tone debate has underlined the fact that we cannot divide the scientific process in a context of discovery and a context of justification. Science is also a matter of people connecting with each other, talking to each other, agreeing and disagreeing, and this has epistemic consequences. Who is allowed to join in the debate has consequences for the knowledge produced. If, for example, the superstars don’t call the shots we can expect established knowledge claims to be challenged more easily. If minority groups feel free to participate then their particular standpoint may have an influence on the kind of knowledge produced. How the conversation proceeds equally has consequences for the knowledge that is produced. A polemical style of debate fits a hypothesis testing approach (remember Popper’s ‘friendly-hostile collaboration’), whereas explorative approaches, or modelling for example, may thrive with less polemical, more constructive kinds of conversation.

Finally, history tells us that there are different models of intellectual conversation, different practices of criticism, different kinds of critics. Parrhesia, to start with, was associated with various practices. There was the democratic parrhesiast, freely giving their opinion as member of a public forum. But parrhesia could also be the guiding principle in the relation between a ruler and his counsellor, or

---

4 The argument that there is no ‘reform community’, which is sometimes heard, could also be seen as a rejection of institutionalisation.
between a master and his pupils. The Cynic was a different sort of *parrhesiast* again: an uncivil outsider, using shock tactics to challenge the norms of society. Polite conversation in the 18\textsuperscript{th} century was also practised in different ways. I mentioned the culture of the Parisian *salon*, where polite conversation was watched over by a *salonnière*. In London on the other hand philosophers and other intellectuals met in coffee houses, and their polite conversation was guided by the rules of the house. And new critical practices are now being created. The kind of ‘data policing’ practised by people like James Heathers in psychology and Elisabeth Bik in biochemistry is, as far as I can tell, a new kind of critical practice, facilitated by new technology. In all these cases power and subjectivity and the relations between them are different. I wonder if there isn’t value in that diversity as such. The debate seems framed by these two poles of *parrhesia* versus politeness, but that framing hides considerable diversity.
Bibliography

Brown, N. (2019, April 18). One of my favourite cartoonists also gets it right on tone policing. https://twitter.com/existentialcoms/status/1118920837405532161 ... [Tweet]. Retrieved April 19, 2019, from @steamtraen website: https://twitter.com/steamtraen/status/1118930465459773440?s=19


Dahlgren, P. (2018, May 13). Peter Dahlgren on Twitter: “@statsepi @VariabilityBlog It’s funny because the research on fallacious reasoning points out that ad hominem is sometimes a legitimate and reasonable argument. A ban on ad hominem is contrary to research on ad hominem, which is even funnier when you think about it. @lakens
https://t.co/74i9DQ8B3t https://t.co/CTe6QFvBp" / Twitter. Retrieved August 30, 2019, from Twitter website:
https://twitter.com/peterdalle/status/995635295331803136

Dahly, D. (2018, May 12). No, we can’t censure people for ad hominem attacks in scientific discourse. Retrieved January 4, 2019, from Darrendahly website:
https://darrendahly.github.io/post/2018-05-12-pnas/

Devezer, B. (2019, March 6). Berna D. on Twitter: “@farid_anvari @IrisVanRooij For my part, I see criticism as destructive, yes, but also an attempt at changing the other party’s opinion (which to me isn’t relevant to scientific discourse). Critical analysis, on the other hand, is something you do either to learn/clarify your own thoughts or to advance &gt;&gt;” / Twitter. Retrieved August 19, 2019, from Twitter website:
https://twitter.com/zerdeve/status/1103392630975545344


Gelman, A. (2016, September 21). What has happened down here is the winds have changed. Retrieved April 6, 2017, from Statistical Modeling, Causal Inference, and Social Science website:
http://andrewgelman.com/2016/09/21/what-has-happened-down-here-is-the-winds-have-changed/


Heemstra, J. (2019, June 26). (3) Jen Heemstra on Twitter: “How did it become part of academic culture to prove how ‘rigorous’ we are by trashing others? How does that advance our field? If we really care about the research, we’ll lead by example in holding ourselves to a high standard and aim our feedback at building up not tearing down.” / Twitter. Retrieved June 27, 2019, from Twitter website: https://twitter.com/jenheemstra/status/1143824625895452672


Jussim, L. (2019, September 2). Lee Jussim on Twitter: “There’s a whole newish SJ lexicon of delegitimization: Grifter, bad take, punching down, bropenscience. Once you have a callous disregard for facts, you are liberated to call anyone you’d like to silence any insulting thing you’d like. Phrenologist, eugenicist, rac/sex-ist... https://t.co/AFiy8sHPYU” / Twitter. Retrieved September 16, 2019, from Twitter website: https://twitter.com/PsychRabble/status/1168610724802961408

Lapan, C. (2019, July 24). Candace Lapan on Twitter: “Interactions like this are why again I will continue to point out the need for #KindScience. To attack and mock people for differences of opinions in science is so unwelcoming and drives people away. I hope we can all find a way to make science kinder.” / Twitter. Retrieved September 16, 2019, from Twitter website: https://twitter.com/CandaceLapan/status/1153794526625980421


Letzter, R. (2016, September 26). We talked to the scientist at the center of a brutal firestorm in the field of psychology. *Business Insider Nederland*. 12


Navarro, D. (2019, August 31). Danielle Navarro on Twitter: “I like this framing a lot. I won’t speak for experimentalists but from a modelling perspective I think we ought to avoid dramatic narratives around ‘crises’. I think it encourages conflict, panic and overreaction, when thoughtful and careful improvement is needed instead. https://t.co/dAXFHZKPbV” / Twitter. Retrieved September 1, 2019, from Twitter website: https://twitter.com/djnavarro/status/1167915934247317504

Nosek, B. (2019, January 15). The feeling of pleasure for a scientific result that came out as desired is usually due to ideology or self-interest, not truth-seeking. That isn’t all bad, but part of intellectual humility is cultivating a sense of interest (at least) for results counter to expectation/desire. [Tweet]. Retrieved January 15, 2019, from @BrianNosek website: https://twitter.com/BrianNosek/status/1085241724207812608


Rooij, I. van. (2019, March 6). Unpopular opinion: We can do without ‘criticism’ in scientific discourse, both the word and the activity. Scientific critique =/=
criticism. [Tweet]. Retrieved March 7, 2019, from @IrisVanRooij website: https://twitter.com/IrisVanRooij/status/1103366445939986432?s=19


Simpson, D. (2019, February 3). And for the love of god do your best to move with humour, humility, kindness, and compassion. Try not to centre yourself as the person who is right, but as someone with intellectual humility who has some questions. [Tweet]. Retrieved February 3, 2019, from @dan_p_simpson website: https://twitter.com/dan_p_simpson/status/1091894345131802625?s=19

Srivastava, S. (2018a, May 13). Tone criticism is a derailment tactic—It changes the subject from the substantive issue to how it is expressed. Allow the established and powerful to define what is acceptable discourse and censure violators, and they will abuse it. [Tweet]. Retrieved May 14, 2018, from @hardsci website: https://twitter.com/hardsci/status/995700997916839936

Stuart, S. (2019, July 26). Shannon Stuart 🌈 on Twitter: “10/ awesome people have addressed much better than I can), one of the things that I find challenging in some academic circles at the moment is the cultivation of a sense of intellectual one-upmanship, with a side of sneer. One of the everyday ways I see this ‘masculinist ethos’” / Twitter. Retrieved September 4, 2019, from Twitter website: https://twitter.com/_shannonstuart/status/1154618969526980608