

Exploring in-group heuristics in peer review panels assessing societal Impact

Dr Gemma Derrick Director Centre for Higher Education Research and Evaluation

Peer review



"The essential principle of peer review is simple to state: it is that judgments about the worth or value of a piece of research should be made by those with demonstrated competence to make such a judgment." The British Academy

- Pillar of academic self governance, The gold-standard in research evaluation
- Benefit of a discussion by peers and experts Democratic deliberation model of evaluation
- Peer review of scientific impact works as there are shared community understandings, interpretations and value of "scientific impact" (Lamont, 2009)
- Peer review outcomes only accepted if the process is perceived to be fair by the research community (Tayler, 2006)
- Resolving differences helps navigate "ambiguous evaluation objects" Derrick (2018) e.g. Interdisciplinary research; Impact and transformative



Peer review: Evaluation by groups



"The dominant definition" (Derrick, 2018)



How do groups work?



- Guided by a negotiated definition of "excellence" (Janis, 1982;
 Levi, 2011; Thaler & Sunstein, 2008; Sunstein & Hastie, 2015)
 - Not always the same achieved through deliberation
- IMPACT IS DIFFERENT ("ambiguous", Derrick (2018))
 - Who is considered an "expert" and who is a "peer"
 - Reaching a common understanding through discussion difficult
 - Evaluators have little prior experience to base decisions
 - Bring in new, different biases and tendencies to evaluation
 - Risk of time poor process no time for experimentation and for things to go wrong
 - Adopt a pragmatic approach to evaluation
 - Potential misuse/use of proxies

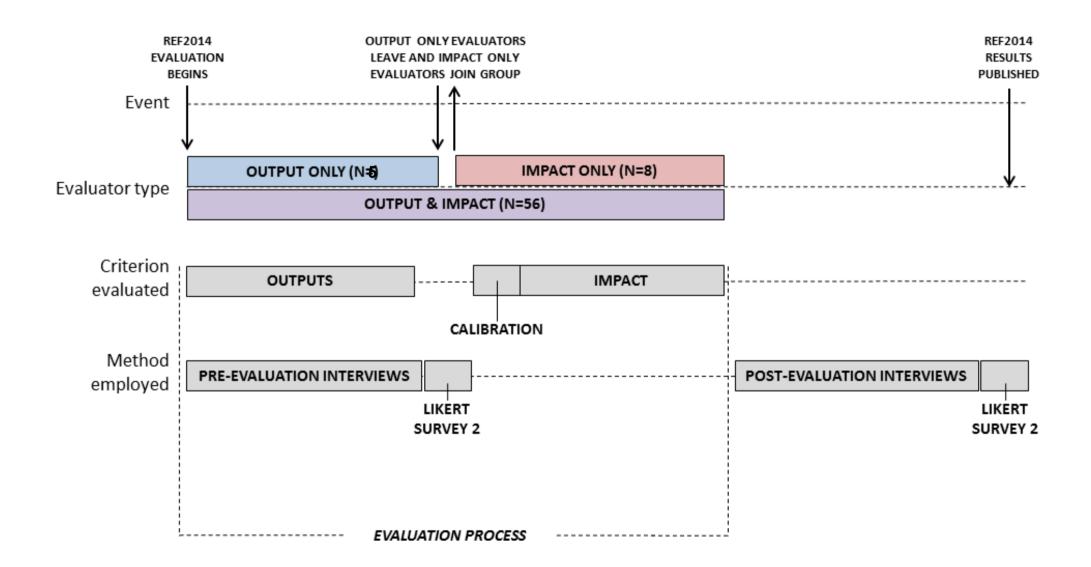
Groupthink



- The consensus is more important than receiving the "correct" decision
- Problem?
 - "defective decision making" Leana (1985)
 - Influences "group cohesiveness" Callaway et al, (1984); Aldag et al 1993)
 - Over-compliance McCauley et al, (1989); Packer (2009)
 - Amplifies influence of interpersonal traits and interpersonal ties aka biases (Riccobono et al (2016)), conscious and unconscious
- Examples
 - Bay of Pigs fiasco, Space shuttle Challenger, Clinical reasoning more recent....and Everything about Trump!!!
- PR legitimacy based on democratic deliberative models of evaluation. Then Groupthink is seen when active moves are made to avoid deliberation as a problem solver (to reach a consensus)

Research design







Group dynamics in peer review

Group psychology tendencies such as identity setting all identified

"The panel, I'd just like to emphasis, you know, we treated the whole exercise as an exercise in collective responsibility, and our panel decided to do things and I'm with the panel on that. I have a personal view but that is not my public view."

Academic (Post-evaluation)

"And so we spent a reasonable amount of time discussing these because it was a way of establishing our principles when we looked at all the others."

Academic (Post-evaluation)

"...we were like a team, suffering together"

Academic (Post-evaluation)

"....I don't know how other panels work, but my panel was quite – perhaps had quite a strong culture shall we say."

User (Post-evaluation)



Groupthink symptoms: Social loafing

Social loafing

- Deferring discussions to others or placing more weight on perceived expertise
- No exact benefit from "loafing" if all evaluators were novices
- Evaluation "Hawks" and "Doves"
 - Dangerous proxy if loafing occurs from dominant voices

"If you get loud people who contribute strongly in their own area they bias the assessment of their own area but then they contribute across the piece that's quite helpful. And other people who are very, very good in their own area but never say anything about anybody else's area."

P1OutImp5(POST)



Groupthink symptoms: Shelving

Shelving

- Avoiding issues not addressing them in evaluation
- Shortcomings of "Impact" are well known (causality, attribution etc.)
- Problem for Impact diverts attention away from deliberation of nature of object (a.k.a what is Impact)
- Allowed panellists to be overly simplistic

"I think most of us feel that although it's a pragmatic way of doing it, it kind of works, it isn't really an accurate reflection of what we've been striving to measure" P1OutImp4(POST).



Groupthink symptoms: Overpragmatism

Over-pragmatism

Avoid deliberation and instead take shortcuts due to time constraints

"...individual impact case studies could have been a bit roughly treated just because of the speed at which decisions had to be made. There would have been no better way to do it than we did it, unless more time were devoted to it"

P3OutImp2(POST)

- Leads to evaluation decisions being "signing off whenever agreed" P5OutImp4 (POST)
- Directly dismisses conflicts that contribute to a robust deliberation, and therefore a more robust evaluation outcome.



Groupthink symptoms: Satisficing

Satisficing

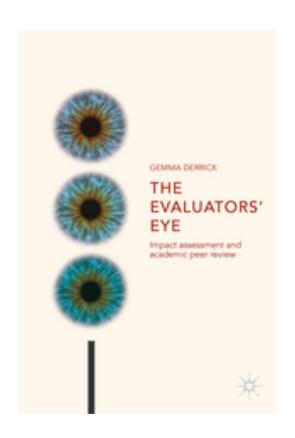
- Suboptimal decision-making strategy (Barge & Gehlbach, 2012)
 - Leads to a decision that is "good enough" and not the best one.
 - Dangerous to legitimacy of peer review as an evaluation tool
- Panels need to weigh up the benefits of obtaining more information versus the cost of continuing a discussion
- REF2014 mechanics (Derrick, 2018) allowed the group to satisfice.

"so we stuck to the script, I have to say" P2OutImp7(POST)

 On many occasions used to re-orientate deliberation if it "drifted away"

Conclusion





- Focus on evaluation practice, and not just outcomes, groupthink behaviours are identified.
- Some have more worrisome influences on the outcome, than others
- Can amplify biases and risk legitimacy of peer review as an evaluation tool for Impact, and similar ambiguous criteria (Derrick, 2018)
- REF2012 Impact now 25% o=so how we evaluate it is now more important

Pre-order on Amazon

Derrick GE (2018) The evaluators' eye: Impact assessment and academic peer review. Palgrave Macmillan.