

Disclosure and Disavowal

Professional Issues in the Scandal in Academia

By Michael James Heron and Pauline Belford.

The Scandal in Academia is a large-scale fictional ethical case study of around 17,000 words and fourteen separate revelations. They are delivered as newspaper extracts from a newspaper reporting on an ongoing crisis at a Scottish educational institution. The scandal case study as presented in its original form comes with only limited commentary on the ethical issues raised, concentrating instead on providing the scenario in isolation. This paper is a companion piece to that case study, discussing the 7th and 8th revelations with reference to the issues raised, the mainstream media, and the formal academic literature. The discussion presented here is not intended to be exhaustive or definitive. It is instead indicative of an approach that could be taken within a formal educational context, and illustrative of the kind of discussions that ideally emerge from the effective use of the material.

Keywords: Ethics; Morality; Professional Issues, Education; Case Study

Categories: K.7.4 [**Professional Ethics**]: Codes of ethics; Codes of good practice; Ethical dilemmas.

Corresponding Author: *Michael James Heron*

Email: *m.j.heron1@ru.ac.uk*

Introduction

The scandal in Academia (Heron and Belford, 2014; Heron and Belford, 2015a; Heron and Belford, 2016) is an extended case study intended for use as a teaching and discussion aid for educational practitioners looking to introduce elements of computer ethics into their curricula. Inspired by Epstein (1997), it is a full-cycle scenario involving many individuals which touches upon the complexity and interrelations of modern computer ethics. It has been trailed and evaluated as a teaching tool by the authors (Heron and Belford, 2015b) but its utility as a general resource is limited without the academic context that supports deeper investigation of the material. It is to address this issue that the authors offer a commentary on the Scandal, with a focus on the 7th and 8th newspaper items presented within.

In these articles, we encounter the mysterious Nemesis, a hacker responsible for leaking numerous emails to the press. In these emails we find evidence of collusion in the matter of the Blackbriar Algorithm, although the extent to which this indicates wrongdoing is a matter for debate. We also find out that it seems Professor Blackbriar has been tweeting angrily from his own personal account, frustrated by the way he has been presented in the newspapers. He feels the university senior management have hung him out to dry. Later the account is locked down because Blackbriar claims it was hacked during the night –

some evidence would suggest though that this was merely the easy way out of an embarrassing drunken Twitter storm.

These incidents raise numerous discussion points for the case study. Within this paper we will discuss issues of data security, data sanitisation, and the extent to which ‘noise’ can be cleaned out of data-sets before meaningful analysis be conduct. We will also discuss the frictionless nature of Twitter and how its unrestrained virality can be both a blessing and a curse when dealing with heated and contentious issues. We’ll discuss the informational asymmetry of whistleblowing, and how in the absence of knowledge of motivation we should be wary of what might be selective and self-serving excerpts from a wider and more nuanced set of information.

All of this has to be understood in the context of the dominant academic precept of ‘publish or perish’. Much of academic funding and reputation is dependent on papers being published and cited – this has a powerful effect on the way in which the results of research are communicated and how sometimes weak evidence can be ‘massaged’ to bring it up to a publishable standard.

As with the previous analysis we have published regarding the Scandal (Heron and Belford, 2014; Heron and Belford, 2015a; Heron and Belford, 2016) we make no claim that this is the only interpretation that can be placed on the material. Readers are invited, indeed – encouraged, to disagree with any and all of the commentary we provide. This is not intended to be a definitive analysis, but an illustrative one.

Publish and Perish

The phrase ‘publish or perish’ codifies the expectations placed upon academic scholars as well as anything for which we might hope. Most within academia have encountered the phrase, and more still in research-intensive institutions have felt its impact. It expresses the common expectation that institutions have of their researchers – publish work, often. It is not enough though to simply publish, it is necessary to publish work in venues that will lead to citation, to name recognition, and to the often ephemerally defined impact (Donovan, 2011). The pressures associated with this have increased in modern years. Nothing happens in a vacuum, and we have already addressed some of the reasons as to why this might be the case in Heron and Belford (2016). Competition over academic jobs has become ever more fraught because universities continue to produce more qualified academics than the system can consume (Powell, 2015). This means that the expectations on those seeking permanent faculty employment end up being continually revised upwards. Uncertainty and instability in employment creates a perversely effective incentive system on those scholars least able to challenge it (Shen, 2015) and this in turn creates a powerful ‘free market effect’ in the academy. If the academy wants high performing academics, it can select for that through evolutionary pressures.

On the surface, leaving aside the human impact of often unreasonable expectations, this might even seem like a good thing. The result after all is ever raising the bar of scholarship so that the next generation is notably better than the last. Evidence suggests though that this has not quite worked in the way that might be hoped. When quantity is incentivised, then other things will suffer as a result. When a research infrastructure emphasises citation, it will produce a culture of citation that is independent of the merit of the work. Heron and Belford (2016) discussed some of the ways in which citation can be manipulated to create the external appearance of impact without any of the concomitant benefit for society. Instead of publishing one truly landmark paper, young academics will often aim for the ‘Minimally Publishable Unit’ (Budd and Stewart, 2015) – the smallest measurable amount of information that would permit publication in an appropriate conference or journal. Four smaller papers can often appear more valuable on an academic resume than one large paper, and almost always serves to skew a citation count upwards as a simple consequence of cross-referencing. The Scandal in Academia itself is an example of that – rather than being one unified monograph of several hundred thousand words it is presented as a series of papers each generating its own ecosystem of citation.

Work may be ‘salami sliced’ so as to appear several times in different publication venues – perhaps with a change in emphasis, or a change in advertised relevance. Again, consider the Scandal in Academia itself – it is likely at the end of its publication life to be repurposed as a book which would count as another academic article even if the new contribution to knowledge is minimal. It’s hard to quantify the extent to which this happens though because the nature of different academic fields prevents meaningful analysis. The number of papers considered ‘enough’ varies wildly from field to field and university to university. While this all may seem harmless, it has an impact on the replicability and reliability of scientific results - statistical significance can be skewed if one single study represents multiple data points.

However, the pressure to publish has a more corrosive impact on the quality of the work. Results are often over-inflated, or over-stressed. Statistical analyses can be ‘massaged’ to give a value for significance that is appropriate for publication – sometimes known as p-hacking (Head, Holman, Lanfear, Kahn, & Jennions, 2015). The relevance of work for a publication venue can similarly be distorted to make the case the work has a deeper merit for an outlet than the original research would imply. Negative results may be recast as positive findings. None of this is academic fraud of the kind discussed in Heron and Belford (2014) – merely a small case of ‘refining’ what was discovered until someone will pay attention.

Consider for example a set of data that yields no statistical significance when viewed through a standard analytical test - a T-Test or a Chi-Squared test for example. By applying ever more sophisticated tools to the same data it becomes increasingly likely that one will reveal a statistically significant result. A cynical peer reviewer might ask 'why didn't you get significance from a less obscure testing regime', but it might also be interpreted as a sign of finding something of genuine worth that was hidden beneath the surface. It's not that the reporting is wrong, or even misleading. It's just that if there is a pressure to publish people can be relied upon to find publishable data if they look with sufficient diligence. We often consider statistical significance with a P value of 0.05 to set a benchmark for 'minimal significance' for publication (Nuzzo, 2015). As David Tumblewood says in the hacked email chain, 'things don't stop being interesting at 0.05'. The nature of significance is often over-stated, and it's entirely possible that a significance of 0.07 is substantive enough upon which to build a case for relevance. After all, the 'gold standard' of significance is that there is only a one in twenty chance the results were derived through random chance. A significance value of 0.07 suggests only one in 14.28 odds the results were derived through chance. Who wouldn't back a horse at those odds, especially if you could look at the horse again from a different angle and speed it up accordingly?

Sometimes the data is massaged in different ways – we've discussed in Heron and Belford (2014) the extent to which outliers can be easily discounted from an analysis to bring the figures in line with the baseline for publication. This is a common, and indeed necessary, part of data curation. However, the fact of it matters less than the motivation – good science should be impersonal and not incentivised to find a particular result. With the Blackbriar Algorithm we're dealing with data points derived at the extremes of what mechanical and digital equipment can gather. It's not unreasonable to remove obvious outliers, but the extent to which they are 'obvious' can be influenced by the external pressures to end up with publishable data. Negative results are incredibly important in science, but they are rarely published in certain fields (Matosin, Frank, Engel, Lum & Newell, 2014) and so researchers are disincentivised away from authenticity in their analyses. If you keep looking you might find something positive, so why not keep looking until you do? The negative result may be more valuable but the positive result is more publishable even if it may end up being misleading.

These are largely internal pressures that come from the publish or perish mind-set – an urge to find the publishable units in the work you do because otherwise you're not going to keep up with a tempo of publication set by those competing for the jobs you want. There are external pressures too, often set at an institutional level. Many academics work within fixed expectations of publishing - either formal parts of their contract of employment, or an informal expression of their local culture. 'Two papers per year in high impact journals' is a common benchmark in research-focused UK universities, even if it's rarely formally expressed. As it is with many norms in a highly mobile workforce like academia, it's sometimes more a diffusing of hearsay than a formal contractual obligation. We carry the expectations of our former workplaces into every new one we enter.

On top of this we can layer formal expectations that influence expectations of publication – papers are what get you academic credibility, and credibility is what finds you collaborators. Collaborators are what you can use to get grant funding, and grant funding is, eventually, how you pay for your salary. Research active academics cannot simply focus on papers – those papers have a role to play in a funding landscape that is intensely competitive. It is cynical to say that universities are only interested in research funding, but it's hard to deny that's how they often behave. More publications equals more credibility which equals more funding. The value of the scholarship often seems to have diminished value in comparison to the fact of its publication.

Consider for example the subtle, and not-so-subtle, hints that Professor Blackbriar receives in the hacked correspondence. 'I don't need to remind you just how important this collaboration is for the university, and for your own research agenda', 'The consultancy fees that ScotOil are offering are enough to fund your department for five years', 'Don't spook the horses'. It's not the case that anything as simple as managerial pressure is being brought to bear here – it comes in a more insidious, more plausibly deniable form. It comes as well-meaning guidance that is looking out for Blackbriar's best interests if he could just shake off his understandable uncertainty about the results. 'As long as you indicate in all publications and correspondence that you're working with processed rather than raw data, I don't see an issue with what David is suggesting', says the chair of the research ethics board. He's right too – there's no academic fraud if you're honest about everything you're doing. However, there's a cynical degree of 'knowing' in advice like this – delving into data sets for replicability happens infrequently (Heron and Belford, 2014) and given the commercially sensitive nature of the data it's not the case that even those directly working on the project could assess the validity of the processing (Heron and Belford, 2016). We see that too in the Blackbriar emails leaked by Nemesis. When most journalists and managers will read the top-line only, they rarely have the time or expertise to consider the implications of a disclaimer.

It is this more than anything else that sets the behaviour of Blackbriar and his colleagues in context. Heron and Belford (2014) discussed the cases of Diederik Stapel (Markowitz & Hancock, 2014), Yoshitaki Fuji (Nomile, 2012) and Dipak Das (Retraction Watch, 2012) as outright cases of fraud and academic misconduct. Here, everyone is telling the truth. They're just not telling the whole truth. They have found a horse and said it was a zebra because if you looked at it at a certain angle, in the right light, you could 'kind of' see some stripes. This can often happen through over-enthusiasm on the part of an amateur – indeed the word 'Zebra' is a kind of medical slang in itself. 'When you hear hoofbeats, think horses and not Zebras'. Sometimes invoking the possibility of a Zebra is more conscious and intentionally manipulative.

It's also important to note here that part of the guidance received by Blackbriar was an exhortation to 'trust your gut'. Widely understood, but rarely explicitly spoken, is the expectation that research is to some extent guided by intuition. We cannot discount here that Blackbriar is a professor of considerable renown and experience in his subject area. While in the end we must always yield to evidence, there is a grey-area where we are still trying to intuit what the evidence is actually saying. In this case, it can be argued that Blackbriar is being pressured into making claims that the evidence cannot reasonably bear – but it could also be argued that this is a necessary part of divining whether this evidence is actually answering the question we think we asked. 'Garbage in, garbage out' is a common phrase in computer science, and it is invoked here by Professor McManus. Perhaps the problem isn't with the analysis – perhaps it genuinely does lie with the data collection. While much of the conversation we see is suggestive of a conscious attempt to mislead it can also quite easily be interpreted as a professional, appropriate discussion on the nature of evidence. Sometimes we may see flaws in results that we cannot necessarily articulate. Experience is a powerful guide to where we may see meaning in aberrations, but it can also be something that calcifies our thinking and makes us resistant to change. Part of the job of an academic is to know the limits of their own epistemology. The 'plea to authority' is often considered a logical fallacy, but to dismiss expertise as irrelevant in interpretation is to succumb to the same fallacy in reverse.

Blowing the Whistle, and Why

One thing that we haven't yet discussed in this paper is the expectation of disclosure that the cited email chain implies. We now live in a world of largely frictionless data transfer – the increasing capacity and speed of modern computers has been more than a difference in scale, it's a difference in kind. Fifty years ago, it would not be possible for Edward Snowden or Chelsea Manning to download gigabytes of sensitive information and make them easily available to journalists and activists (Heron, 2016) – the act of removing that amount of data from a secured facility would have involved physical media and a supply of wheelbarrows. Now, it's a matter of a few minutes to transmit vast amounts of information across the globe. This has had a remarkable impact on public discourse and the expectations we have of transparency for our governments. 'Information wants to be free' is a common slogan of free speech and free information advocates, and when information becomes free it is very difficult to chain it. Consider for example the encrypted torrent file Wikileaks distributes as insurance (Cammaerts, 2013) – in the event anything happens to their central repositories all they need to do is release a password and everything contained within will become public knowledge.

Along with the difficulty in containing information once it is revealed comes increased difficulties of preventing its revelation in the first place. Academics and public figures are often the targets of malicious, or merely inquisitive, attempts to access their private data – primarily that stored on cloud services, or on remotely accessible servers. Unintended disclosure is not a new problem - journalists of the less reputable outlets have been known to trawl through the garbage of celebrities in the hope of finding reportable information. The whole espionage business is conducted in an ongoing exercise of escalating revelation – using information to lever open access to other pieces of information. However, the ease at which this can now be done remotely, and the quantity of data that can be obtained, has changed the texture of privacy online.

Consider for example the consequences of a badly chosen and easily guessed password. If used too liberally between multiple services it might unlock access to years of email, social media accounts, chat logs, photo repositories, and more. Many of the things that might be revealed could be incriminatory, or open access to other parties involved in discussion. We all use many digital services on a daily basis, and few of us know for sure how diligently these services are guarding our data. The best services are resistant to hacks and leaks, storing sensitive data in an encrypted format and ensuring it does not yield itself to brute-force password attacks. With the growth of encryption came a corresponding increase in the sophistication of hacking. To counteract the time taken to compute password hashes, precomputed rainbow tables became a common tool of those looking to force access where they were not invited. That in turn led to the practice of ‘salting’ a database so that rainbow tables became much less effective. The war between those that want access to data and those that wish to prevent it is an escalating one, with new fronts opening on a daily basis. However, institutions and services cannot necessarily be in the vanguard of the war – it is costly, requires expertise, and is often highly contextual and situational. As a result, we are all more vulnerable to unwanted disclosure than we might like. Equifax was recently responsible for a massive leak of incredibly sensitive financial data – around 150m records. In 2013 and 2014, Yahoo was responsible for leaks that impacted one and a half billion users in total. A hack at AdultFriendFinder in 2014 revealed the sexual proclivities of over 400m users. In 2015, hackers stole gigabytes of user data from the site Ashley Madison, which sold itself as a venue for married adults to have a discreet affair. The true impact and severity of these leaks varies from case to case, but it’s clear that we cannot trust other parties to look after our data for us.

The impact that these kind of disclosures can have can be seen in the articles we've been referencing through this paper. It is often not the case that what we have written is sensitive information, but rather information that has a certain expectation of the context of its audience. Within a technical publication we might reference jargon, common tools and techniques, and make reference to technology and concepts that could not be expected of a 'lay' reader. That's entirely appropriate because academic papers are intended to be read by a specialist audience, in the same way that procedural motions in a court are intended for a legally literate audience rather than the populace as a whole. We would not use such specific terminology when describing the work to friends and relatives with whom we cannot assume common ground in understanding. There is a context of audience that we all implicitly function within. However, even in more informal communications there is a form of social context that textures and informs the way in which we discuss topics. This context is rarely explicitly referenced when leaked emails and such are published for the unguarded and salacious gossip they contain. Communication is not simply statements delivered in order – it's a collaborative exercise in building understanding. Both the source and recipient of information are engaged in a collective task of interpreting meaning, and the literal definition of words and terms by themselves may not fully illuminate that intended meaning.

Consider for example the simplest of these contexts – familiarity. In-jokes and shared experiences can radically change what words mean and how they should be interpreted. A shared bond between people can ensure that jokes are understood as such, and sarcasm is picked up for what it is. In taking a random email from a random inbox, we are ignorant of such nuance and it's impossible to codify it in a form that genuinely captures the often heavy weight of contextual subtlety. 'You could always fake the data', might be intended as a joke when sent to the designated recipient. The recipient might understand that instantly. No serious intent or suggestion of fraud is intended. However, when read uncharitably by an otherwise dispassionate independent observer it looks an awful lot like an exhortation to misconduct. More than this, attempts to explain the context almost always come across as insincere justifications – when our statements are distributed wider than our intended audience, we are put in an unwinnable situation because the attempt to justify seems defensive. Consider for example Professor Barry Spurr (Davis, 2014) who resigned from his position as a professor at the University of Sydney after a leak of racist and sexist emails. His defence that the emails were a 'whimsical word game with colleagues and friends' convinced few.

Whether that is the case with the Blackbriar emails, we will never know. Is the hectoring tone of publication pressure a jovial reference to their shared contempt for the philosophy of publish or perish? Perhaps ScotOil is something of a joke between the participants in the conversation, and so ‘I don’t need to remind you just how important this collaboration is for the university’ might be a caustically arch statement that carries the opposite meaning of what the words say. The use of the phrase ‘representative historical data samples’ is in quotation marks. Is that an in-joke? Does it refer to a draft institutional policy? We don’t know and can’t know, and as such we should be wary of reading too much intent into such emails. It’s common for the misquoted to claim ‘out of context’ for statements – that’s not true in this case, because the full surrounding context of the remarks is available. However, context can incorporate more than the words that precede and follow the indicated statements.

With our mysterious hacker too, we must appreciate that there may be a hidden agenda at work. The nature of an informational leak is that it is asymmetric – we cannot know how much information has been leaked, or what impact the order in which it is released is intended to have. The very nature of revealing hidden information is shrouded in secrecy because we do not know what remains hidden at the end of the process. We do not know if this is the full set of information, or whether later emails would contradict or perhaps further contextualise the earlier ones. Could we honestly say, in all fairness, that this leak even related to the project in question? We are led to believe that it is the Blackbriar algorithm that is being discussed by the dates and the juxtaposition of the emails with the ongoing coverage. Perhaps this is Blackbriar commenting on the work of another colleague. We simply don’t know because we have been given cherry-picked emails that are devoid of their full, unambiguous situation in a conversational record. They certainly seem suggestive, but that does not constitute evidence in its proper sense. Consider for example the Climategate (Leiserowitz, Maibach, Roser-Renouf, Smith & Dawson, 2013) emails. These had been obtained through the hacking of a server at the Climatic Research Unit at the University of East Anglia. They were heavily publicised in the climate change denialist press as evidence of a sustained conspiracy with regards to climate change research. Certain emails could be interpreted in such a way as to suggest conflation of conjecture and conclusion, manipulation of results, and scientific trickery of the kind outlined in the Blackbriar case we have been discussing. Eight independent local and national committees investigated the emails, and found no evidence of fraud or scientific misconduct. The informality of communication and the coining of internally consistent jargon permitted skeptics to read more into the communication than a fairer, more realistic reading would permit. The Guardian, reporting on the scandal, noted that there was much to suggest that the emails released had been filtered to give undue prominence to a handful of emails that could be most easily interpreted in an unfavourable light. Terms such as ‘Mike’s Nature trick’ were used to describe methodological techniques that could meaningfully connect together two different kind of data sets (Skrydstrup, 2013). Within the expected audience of the emails, the term was understood to refer to a broadly understood and accepted analytical practice. Outside of that context, it took on a much more sinister tone.

This shines a light on why whistleblowing is not necessarily an ethically pure phenomenon, especially when it is being done by an anonymous party. There are many reasons why disclosure of identity would not be desirable when reporting on wrong-doing and it would be churlish to assume all those that maintain their distance from disclosure are doing so for dishonest reasons. However, the fact the action is reasonable does not inoculate it from its troublesome aspects. An anonymous party may have no agenda, or they may have a very pointed agenda. They may have revealed all the information, some of the information, or some highly targeted subset of the information. They may have released the information unedited and uncurated, or they may have edited key elements of the data to further their own agenda. They may have unfairly focused attention on one participant in a conversation while shielding others. Within the Scandal in Academia we do not know who Nemesis is – at least at this point. The agenda that they harbour is something we need to know before we can assess the information in its fullest light.

Data security is important within academia because we are often operating under formal requirements for data management. We may be beholden to funding partners to protect commercial interests. We may be obligated by a university's research policy to make available full data-sets unedited and uncurated. We may have an ethical duty of care to ensure the anonymity of research subjects and participants, and we may have a similar ethical duty to ensure the protection of their contributions. We will also almost certainly be under a legal responsibility to handle data in line with existing legislation – the Data Protection Act and the Freedom of Information Act in the UK being two obvious examples. And yet, we are often pulled in multiple directions – we need to lock down information for protection while opening it up for collaboration. University architectures are often insufficient for dealing with the flexibility and variety of data needs and academics often resort to the use of external tools – dropboxes, torrents, servers outside of university firewalls, and more. This is often as much for necessity as it is for convenience, but every service that we use adds another point of weakness for those looking to gain access to information over which we would prefer to have control. Often the disclosure of the data would present no serious or significant problems – it's often hard enough to get people to pay attention to research even when you're screaming about it to everyone you can at every opportunity. Occasionally though, even the most innocuous of data can be compromising if wielded by someone intending on doing damage and unconcerned if the damage may only be temporary. The initial attack may be front page news after all. The retraction is most likely to be a small notice buried in the back pages of a later edition. As such, what we see here in the Scandal in Academia is perhaps indicative of an ethical lapse from all participants in the management structure. It might just as easily be an out of context attack from an aggrieved party with their own axe to grind. Whistleblowing is a double edged sword for everyone involved.

Tantrums and Temptation in Social Media

It's not uncommon in modern times for those at the centre of a media storm to lash out on social media platforms. As unwise and unhelpful as it usually is such actions permit a degree of control of the message that is otherwise unavailable to an affected party. The frictionless nature of the digital age that we discussed above also extends to social media platforms – a single tweet can become a global discussion point in the time taken for one

person to fly from New York to South Africa (Ronson, 2015). Most social media platforms are built upon a principle of sharing, and designed to make sharing as easy as possible. Facebook has its share button. Twitter has its retweets. At the press of a button a signal can be boosted by a reader, and then boosted by their readers, and so on. Much of the battle for attention online is fought in a contentious marketplace of trending topics where there are no clear strategies for ensuring who the audience will be, or how the topic will be interpreted. The ‘virality’ of content is a marketable trait of considerable value when wielded wisely (Hansen, Arvidsson, Nielsen, Colleoni, & Etter, 2011). However, the Internet has a tendency to flatten nuance and dilute subtlety, and as we discussed above even determining meaning is a complex act of negotiation.

With this in mind, it’s common for aggrieved parties to wish to state their side of the story. Such attempts only rarely go well – often by the time a situation has gotten to this point it has already been spun into a dominant narrative by those participating in conversation. Nobody on social media gets an opportunity to ‘control the story’ – at best they can try to nudge it onto a more agreeable trajectory. The temptation though can be hard to resist when one otherwise feels themselves silenced or censored by more traditional media outlets. The online community is febrile and amplifies the extremes of conversation through anonymity (Lee, 2007), deindividuation (Williams & Guerra, 2007), and simple weight of numbers. These elements come together without engineering, but can also be corralled and directed as a specific unit of highly tailored harassment (Heron, Belford and Goker, 2014). It is estimated that Twitter has an active user base of around 328m per quarter. With such numbers, even if we consider a particular event to be a ‘million to one’ occurrence it’s still happening over 300 times every three months. The sheer mass of activity on Twitter is its own system for amplification, and often not in positive directions. A recent study (Demos, 2016) suggested that messages that could be legitimately defined as misogynistic came in at a rate of 6.6 abusive tweets per minute. The pattern of abuse too is not uniform, and overwhelmingly directed at particular target demographics. With the mass of aggression and unpleasantness online, expecting it to be a fertile ground for reasonable exploration of a nuanced situation is optimistic at best. As a personal anecdote, one of the authors of this paper was recently branded as ‘ableist and misogynist’ by a Facebook user for giving prominence to issues of colour blindness on a website devoted to the accessibility of tabletop games (<http://meeplelikeus.co.uk>). Social media removes the friction that exists in real world interactions, and permits both a kind of ‘drive by’ system of aggression and removes almost all of the social context that would moderate face to face discussions. Such a comment in the real world would be possible to discuss and dissect, and would likely be delivered in a way less likely to instantly create friction. We are all more circumspect in our opinions when there are real world social consequences that are linked to their expression. However, the nature of Facebook permits for such perceived ‘truth bombs’ to be dropped into a public space, detonated, and then for the source of the antagonism to disappear behind a shield of blocking and muting that prevents follow-up and meaningful discussion.

While many platforms permit for injudicious comments to be deleted or edited, that doesn’t allow for commentators to undo the damage they may have done to themselves by

interjecting an opinion into a fraught discussion space. Tweets and status updates may be screenshotted, and online services exist that archive tweets as they are made so that the fact of their removal or editing can be given undue prominence. Sometimes the act of recanting, or clarifying, will do little more than add fuel to a burning fire. Individuals should enter into the arena of social media carefully because it is unforgiving of mistakes and missteps. We can see this when Professor Blackbriar, frustrated by what he deems to be media complicity in a stitch-up, lets a momentary lapse in judgement potentially steer the course of the rest of his life. Presumably drunk, given the content of several tweets, he lashes out at Dungen senior management accusing them of complicity and pressure with regards to publication of dubious data. The party previously seen as ‘wronged’ was likewise implicated in a conspiracy to mislead shareholders – ScotOil was alleged to have had a more active role than we have previously seen. It was claimed the Senate and Court of the university were directly involved in everything that was going on, and even his lawyers get something of a broadside when he acts against their explicit advice to ‘keep quiet’. Blackbriar clearly believes that he is being slandered and that social media is the weapon he can use to clear his good name. In the cold light of day, six hours after the last tweet on the topic, Blackbriar claims that he has been hacked and his account is locked.

We have no way of knowing the truth of this, but it has to be said it rings hollow – and not just because it’s so common a defence against social media misconduct that it has become clichéd. This may perhaps be the only way Blackbriar can regain control of the situation – through actively sacrificing his own freedom of speech because of the dangers it carries with it. Certainly within the UK, the fact that a comment is made on social media does not inoculate its maker from legal sanction. UK blogger Jack Monroe won a case against Mail Online columnist Katie Hopkins over defamatory tweets. The judge found that there has been “serious harm” to Monroe’s reputation due to Hopkins’ comments. American actor James Woods won a legal action against a dead Twitter user who had accused him of cocaine use. Sally Bercow arrived at an out-of-court settlement in a defamation case raised by Lord McAlpine. The judge ruled that her tweets were “seriously defamatory” and legally indefensible. The conversational nature of the medium, and the heavy role that interpretation plays in meaning, is not in itself a legal defence when circumstances go beyond that which can be settled with words alone. This is still largely uncharted territory as far as the courts are concerned. These early indications suggest that there is a weight of consequence that must be taken into account when potentially libellous statements are made in a social media forum.

With that in mind, it is perhaps not surprising that a measure of cold reality seeps in when someone has made a mess of a situation via social media interaction. Occasionally this can genuinely be attributed to hacking, but often the perceived anonymity of hackers serves merely as a convenient shield behind which wrong-doers can hide. After his account posted a photo showing ‘a real lamping’ of 100 foxes, Vinnie Jones claimed “this is a hack ive never seen this pic in my life and did NOT tweet it is a hack !!!”. Nadia Gustavo was a South African Zumba trainer who’s account posted “Friday 20, 2016, three black kaffas were finally found guilty of armed robbery in Queenstown regional court...”. On losing her job she said “I have just lost my job at Virgin Active for something I did not do”.

Amanda Bynes claimed she was hacked in a twitter spat with rapper Kid Cudi. Anthony Weiner initially blamed a mistakenly posted photograph of his penis on hackers, although he later recanted and apologised. Sometimes it's not hackers, but unnamed and unidentified assistants or staffers. Ted Cruz 'liked' a pornographic video on Twitter, and claimed that it was the result of a staffing issue. It's sometimes not even as a result of aggression, embarrassment or unpleasantness online. Alicia Keyes claimed she had been hacked after an innocuous tweet was posted from her iPhone – not a problem except she had just been made Creative Director for Blackberry. Sometimes the issue is merely that you've inadvertently outed yourself as being a supporter of an inconvenient brand.

We cannot be sure of the truth of such claims without a full investigation being conducted and the results being made available. It is though simultaneously plausible and implausible. Supposition and gut feeling must substitute for genuine insight into the real state of affairs absent sufficient attention directed to the claim by parties able to uncover the truth and make it known.

In the case of the Scandal in Academia, not only are we unaware of the truth we know there is a motivated hacker working in the background to implicate major figures in the scenario. Claims of being hacked are often difficult to believe because of the convenience of their timing – they act as a kind of real-life deus ex machina that leaves people unconvinced. It is perhaps telling that immediately in the aftermath of the twitter comments our anonymous hacker revealed an email that substantiated many of the claims the professor's account had been making. Similarly, an anonymous insider at the company has given reasons to suspect there may be an element of truth to the accusations made against ScotOil itself. However, other emails leaked by Nemesis serve to act against the revelations made – contradicting the Professor's perhaps self-serving interpretation of events by releasing emails that put his claims under a light of scrutiny. If he had been hacked by Nemesis, it seems unlikely the tweets would have run counter to the hacker's agenda. If he had been hacked by another anonymous hacker, how would they have known enough about the situation to make meaningful allegations regarding privileged and private information? On the other hand, perhaps there is an element of disinformation here – it doesn't do a hacker much good to make it obvious they're behind a hack, certainly not if their agenda requires the implication of a third party.

The Post-Truth Era

It's perhaps telling too that in the correspondence between Blackbriar and Sharon McAlpine we see the emergence of an argument that would later find prominence within the Trump candidacy for president. Blackbriar claims in an email, 'The criteria make no sense without access to the full data sets and would actually be misleading'. Eric Trump, in explanation of why his father would not release his tax returns, said 'You would have a bunch of people who know nothing about taxes trying to look through and trying to come up with assumptions on something they know nothing about. It would be foolish to do'.

Whether this is a fair reason or not is up to each of us to decide as individuals. However, there is genuine danger that comes from non-experts reading too much into specialist data.

Much of the climate change discussion, as we have outlined above, is very technical and couched in sophisticated internal jargon and convention. The sheer complexity of data can be a compelling reason as to why it should be ignored if one is pushing an advocacy agenda. Evidence denial (Lewandowsky, Oberauer, & Gignac, 2013) is a major reason behind the growth of pseudoscientific views in modern discourse. Having access to the data permits those that know enough to misinterpret it (knowingly or otherwise) to conveniently ‘contextualise’ it for a non-specialist audience. Most of us must rely on the analysis of experts in complex fields. We rely on doctors to interpret medical research. We rely on lawyers to interpret case law. We rely on teachers to interpret pedagogic theory. Even experts in one field cannot be relied upon to understand the context of another – those that work in health research cannot extrapolate the conventions that are germane to pedagogic research, and vice versa.

We live in an age of decreasing trust of expertise and with this comes a growth in a mindset that believes facts are subject to interpretation. It is sometimes said we now live in a post-truth era. Claims such as these are part of the problem, simultaneously attempting to elevate one viewpoint as unquestionable truth and the other as ‘post-truth’ wish-fulfilment. When applied to raw, unquestioned data this can be a fair comment – one is one, and that’s not up for debate. When applied to complex, multivariate scenarios this is something more likely to alienate people than convince them. As scientific advocacy becomes increasingly politicized, we must accept the possibility that there are multiple interpretations for data and that an expert presents us with the one they find most credible. Ideally this is the one most compatible with the evidence, but it would be foolish to believe that sometimes it’s the interpretation that is most compatible with an expert’s politics or moral framework. It is sentiments such as these, and the occasional obvious example of experts or pundits pushing a biased agenda, that underlie claims such as that of Michael Gove when he said ‘Britain has had enough of experts’. This is a dangerous position for society to be in – when those most able to offer a clear and realistic lens onto data are held in the least esteem by those responsible for acting upon their advice.

That perhaps represents one of the most dangerous outcomes of the publish or perish mindset – that we’re all so busy engineering the circumstances of publication that we don’t have time to worry about the trust we may be eroding as a result.

Conclusion

One of the key measures used to analyse the truth of a scientific principle is how well it permits us to predict the future. A meaningful theory is falsifiable, and makes concrete claims about what will happen in the future should the theory be correct. While such things are not appropriate to claim in a series of papers on computer ethics, it’s notable that in several respects many of the things that happened in the Scandal in Academia (Heron and Belford, 2014) have had their own echoes in modern discourse. There is nothing especially admirable in this since the Scandal itself is drawn from many real life incidents and as such it has only the predictive capabilities of the past itself. However, it does show just how relevant the topic of computer ethics education is and remains. As time goes by,

we need formal systems to contextualise this topic in greater quantities. The world is complicated, and information technology is accelerating that complexity. Things don't get simpler from here onwards.

We are still largely in the dark as to where the real blame lies for what happened at the University of Dungen. As is often the case, the more we know the less sure we can be. The truth is that we are unlikely to ever be able to point to a single individual and say 'Arrest them, officer' – real world complexity has the inevitable effect of smudging over simple, clear judgements. Blackbriar has clearly done wrong things, but he's done them under pressure. Senior management in the university have been responsible for that pressure, but they also have the survival of a whole institution to concern them. There are real question marks over the accuracy of the data used in the algorithm, but we have equal question marks over much published research given the biases the publish or perish model introduces into academic discourse. We might like people to be stalwart paladins, standing true against an encroaching tide of compromise. We draw our lessons about acceptable behaviour from the social context in which we find ourselves, and academic research rarely presents us with the clear, unambiguous answers we would like. As such, we are often compelled to resolve those ambiguities in whatever ways we can – that can be difficult to explain to an audience that has not been immersed in the complexities we have faced. No research plan survives contact with the enemy, if you will pardon the paraphrase. Anything publishable is likely to be battle-scarred because the academic landscape itself is a war between our competing impulses.

As time goes by, we are likely to increasingly be led to the conclusion that doing right is impossible within the complex social environments we find ourselves. Doing the right thing requires an understanding of what that is, and that's difficult when reasonable people can disagree on how we even judge what might be right. That doesn't inoculate us from our responsibilities as professionals – it is our job to work out what's right and act in accordance with it. We won't always be correct, but simply contemplating the issues is an invaluable first step in acting in accordance with the occasionally contradictory edicts of our ethics.

As outlined in the introduction, we make no claim that this is the definitive analysis of the two indicated newspaper items. We seek only to offer a lens through which the scandal in academia can be contextualised within its broader context. We seek to demonstrate why each of the individual articles opens up wider and deeper discussions of the issues of modern ethics and the factors that influence them. We hope that this analysis of the scandal helps inform educators looking to use the case study within their own courses.

References

Budd, J. M., & Stewart, K. N. (2015). Is There Such a Thing as "Least Publishable Unit"? An Empirical Investigation. *LIBRES: Library and Information Science Research Electronic Journal*, 25(2), 78.

Cammaerts, B. (2013). Networked Resistance: the case of WikiLeaks. *Journal of Computer - Mediated Communication*, 18(4), 420-436.

Davis, M. W. (2014). The persecution of Barry Spurr. *Quadrant*, 58(12), 20.

Demos (2016). The Use of Misogynistic Terms on Twitter. [Available online at <https://www.demos.co.uk/wp-content/uploads/2016/05/Misogyny-online.pdf>]

Donovan, C. (2011). State of the art in assessing research impact: introduction to a special issue. *Research Evaluation*, 20(3), 175-179.

Epstein, R. G. (1997). The case of the killer robot: stories about the professional, ethical, and societal dimensions of computing. New York/Chichester: Wiley.

Hansen, L. K., Arvidsson, A., Nielsen, F. Å., Colleoni, E., & Etter, M. (2011). Good friends, bad news-affect and virality in twitter. *Future information technology*, 34-43.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS biology*, 13(3), e1002106.

Heron, M. J. (2016). Ethics in computer science. In Laplante, P.A. (ed.) *Encyclopedia of computer science and technology*. 2nd ed. Abingdon: CRC Press.

Heron, M. J., & Belford, P. (2014). Ethics in context: a scandal in academia. *ACM SIGCAS Computers and Society*, 44(2), 20-51.

Heron, M. J., & Belford, P. (2015a). Power and perception in the scandal in academia. *ACM SIGCAS computers and society*, 45(2), 11-19.

Heron, M. J., & Belford, P. H. (2015b). A practitioner reflection on teaching computer ethics with case studies and psychology. *Brookes eJournal of Learning and Teaching*, 8 (3). [Available online at <http://bejlt.brookes.ac.uk/paper/a-practitioner-reflection-on-teaching-computer-ethics-with-case-studies-and-psychology/>]

Heron, M. J., & Belford, P. (2016). The quantified university: an analysis of the Scandal in Academia. *ACM SIGCAS Computers and Society*, 46(3), 28-44.

Heron, M. J., Belford, P., & Goker, A. (2014). Sexism in the circuitry: female participation in male-dominated popular computer culture. *ACM SIGCAS Computers and Society*, 44(4), 18-29.

Lee, E. J. (2007). Deindividuation Effects on Group Polarization in Computer - Mediated Communication: The Role of Group Identification, Public - Self - Awareness, and Perceived Argument Quality. *Journal of Communication*, 57(2), 385-403

Leiserowitz, A. A., Maibach, E. W., Roser-Renouf, C., Smith, N., & Dawson, E. (2013). Climategate, public opinion, and the loss of trust. *American behavioral scientist*, 57(6), 818-837.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—therefore, (climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological science*, 24(5), 622-633.

Markowitz, D. M., & Hancock, J. T. (2014). Linguistic traces of a scientific fraud: The case of Diederik Stapel. *PloS one*, 9(8), e105937.

Matosin, N., Frank, E., Engel, M., Lum, J. S., & Newell, K. A. (2014). Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture.

Normile, D. (2012). A New Record for Retractions?. *Science Insider* (American Association for the Advancement of Science), 2.

Nuzzo, R. (2014). Statistical errors: P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume. *Nature*, 506(7487), 150-153.

Powell, K. (2015). The future of the postdoc. *Nature*, 520(7546), 144.

Retraction Watch (2012). UConn Resvatro research Dipak Das fingered in sweeping misconduct case. Retraction Watch. [Available online from <http://retractionwatch.com/2012/01/11/uconn-resveratroresearcher-dipak-das-fingered-in-sweeping-misconductcase/>]

Ronson, J. (2015). How one stupid tweet blew up Justine Sacco's life. *New York Times*, 12.

Shen, H. (2015). Employee benefits: plight of the postdoc. *Nature*, 525(7568), 279-281.

Skrydstrup, M. (2013). Tricked or troubled natures?: How to make sense of "climategate". *Environmental science & policy*, 28, 92-99.

Williams, K. R., & Guerra, N. G. (2007). Prevalence and predictors of internet bullying. *Journal of Adolescent Health*, 41(6), S14-S21.