JOURNAL OF CONTROVERSIAL IDEAS

Article

All Signal, No Virtue: How Trigger Warnings and Other Ineffective Pedagogical Practices Spread

Renaud-Philippe Garner [1,*] and Nazaneen E. Jamil [1]

[1] Department of Economics, Philosophy and Political Science, UBCO, 1147 Research Road, Kelowna, BC V1V 1V7, Canada

* Corresponding author: r_p.garner@ubc.ca

Submitted: 5 June 2024, accepted: 27 February 2026, published: 27 April 2026

Abstract: This paper begins with a case study and then makes a broader argument about how the proper exercise of the intellectual virtues is undermined by failures of character and institutional incentives. Our topic is the rise and spread of trigger warnings as a pedagogical tool. In part I, we define them and explain how they spread. In part II, we review the justifications for trigger warnings. In part III, we review the empirical evidence and show how it undermines these justifications. In part IV, we make a broader argument that draws on Aristotle and MacIntyre. Given that there never was any good evidence that trigger warnings work, why are they so ubiquitous? We argue that their adoption and use is best explained by a lack of prudence, which is explained by two other failures. On the one hand, the unwillingness to speak up is due to a failure of character. Pedagogues can read the evidence, but they are unwilling to speak up when doing so is costly and requires courage. On the other hand, educational institutions do not favour virtue because professional success is often at odds with the excellence that is internal to teaching.

Keywords: trigger warnings; education; prudence; conformism; institutions

How to cite: Garner, R.-P.; Jamil, N. E. All Signal, No Virtue: How Trigger Warnings and Other Ineffective Pedagogical Practices Spread. *Journal of Controversial Ideas* **2026**, *6*(1), 5; doi:10.63466/jci06010005.

©2026 Copyright by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.

> We are good, perhaps a little too good, but we are also a little stupid; and it is this mixture of goodness and stupidity which lies at the root of our troubles. –Karl Popper

1. Part I: Trigger Warnings, What and Why

Trigger warnings are a distinct and recent form of warning (Bridgland et al., 2023). Some warnings advise that content is for a mature audience. An X-rated film signals content (ex. drugs, sex, violence) deemed ill-suited for immature viewers. Other warnings advise people that the content is inappropriate for a given context. The expression 'Not Safe for Work' (NSFW) captures the idea that the acceptability of content depends on context.

The ideal trigger warnings have three features. First, they are addressed to a group considered to be *vulnerable*. The textbook case is someone suffering from post-traumatic stress disorder. As they became more widespread, they seemed to address more and more mental health struggles. Second, they alert the vulnerable that the upcoming content is (more) likely to trigger them and cause psychological distress. For this reason, they more frequently precede descriptions of rape than of roadbuilding in Ancient Rome. Third, trigger warnings are meant to help the vulnerable; there is a presumption that the vulnerable are better off when they are used.

In practice, trigger warnings look like other warnings. "Warning: the following contains descriptions of domestic violence" resembles a warning such as "Warning: there are sharks in these waters". Yet, they do not rely on the same logic. Predators like sharks are a threat to all humans, and it is therefore helpful for any human to know if predators are nearby. In the first case, we assume that the descriptions of domestic violence represent a particular threat to the mental well-being or health of certain people. And if descriptions of domestic violence only threaten the mental well-being of some, they resemble warnings about allergens: they are not addressed to humans qua humans, but are intended to help a subset who are vulnerable.

In sum, trigger warnings signal that upcoming material represents a greater risk of emotional distress for those whose past suffering matches the content. A typical example would be "Warning: this contains descriptions of sexual assault" or "Our next class will discuss intimate partner violence and may be distressing". However, warnings like "Viewer discretion is advised" or "This content is not safe for work" are not trigger warnings.

Some might wonder how trigger warnings differ from banal announcements. A history teacher might mention that the next class is on the Cambodian killing fields. Is this a trigger warning?

Two points are worth underlining. First, there is nothing inherently therapeutic about announcing something. Good orators use signposting because it is helpful to announce what is to come, not because it is therapeutic or deeply benevolent. Second, we should avoid pathologising ordinary emotions. The miserable and cruel deaths of two million Cambodians under the Communist Party of Kampuchea should move us. However, our negative emotions and experiences do not all amount to trauma.

We briefly touch upon the prevalence of trauma. It is easy to find the claim, in writing or public speeches, that "everyone has trauma". While there is a grain of truth to this claim, it is deeply misleading and unhelpful. We must distinguish between a) exposure to a traumatic event and b) whether or not this exposure has led to the development of post-traumatic stress disorder (PTSD).[1] Drawing on the *National Epidemiological Survey on Alcohol and Related Conditions* (NESARC-III), the National Center for PTSD reports that the lifetime prevalence for PTSD among American civilians is 6% and 7% for veterans (Schnurr, 2025). Another study found that while exposure to canonically traumatic events is quite high—nearly 90%—the prevalence of PTSD over the short, medium and long term is shy of 10% (Kilpatrick et al., 2013: 537). Further, a systematic literature review found that the lifetime prevalence of PTSD was 3.4 to 8.0% among civilians and 7.7. to 13.4% among veterans (Schein et al., 2021: 2158). These converging results show that nearly all of us have experienced or will experience something traumatic, but very few of us are traumatised by such events. We do not all have 'traumas' in the sense of having PTSD. Therefore, it is crucial to distinguish between the prevalence of traumatic events or experiences and the prevalence of post-traumatic stress disorder.

---

[1] The nature of trauma is debated: whether trauma requires an objective component (a list of trauma-causing events) or whether the subjective component, the memory of the event, is sufficient. Regardless, our point is that the prevalence of trauma is seriously overstated. For the above debate, see Rubin (2008).

Still, the crux of the issue remains unchanged. The use of trigger warnings relies on the same reasoning: if we wish to help those who have been traumatised and are psychologically vulnerable, then we need to warn them when they are likely to be triggered. We can disagree on who is vulnerable, or what makes them vulnerable, but this does not change the reasoning behind the use of trigger warnings.

Online communities discussing sexual violence seem to have invented trigger warnings in the late 1990s (Bridgland et al., 2023: 2; Manne, 2015; Mannix, 2022). Quickly, they spread. By 2014, the student senate of the University of California, Santa Barbara voted a motion to urge instructors to include them in their syllabi (Byron, 2017: 116). Their spread appears to result from activism and imitation. There is no evidence that their widespread adoption followed in the wake of breakthrough research or a landmark study. In part III, we see that research emerged only years after advocates pushed for adopting this new practice.

2. Part II: Why Use Them?

This section reviews the arguments for trigger warnings published in academic journals or broadsheets. The teleological ones claim that they help us achieve some desirable result while the non-teleological ones focus on how they treat people, on intrinsic rather than instrumental value.

Teleological arguments divide into harm reduction and promoting the good. We can identify two variants of harm reduction: *avoidance* and *coping*. Arguments based on avoidance claim that trigger warnings benefit students by helping them avoid distressing material. Interestingly enough, it is rarely endorsed by advocates of trigger warnings and seems motivated by related therapeutic strategies, like the use of therapeutic animals or safe spaces that allow students to exit class or a presentation and find refuge (Shulevitz, 2015). While critics have objected that a quality education does not allow students to avoid difficult topics (Lukianoff & Haidt, 2015), it is hard to find explicit endorsement of the *avoidance argument* (Dickman-Burnett & Geaman, 2019).[2]

The second variant of the harm reduction argument is coping. Upon encountering a trigger warning, traumatised students know that the time has come to use coping techniques. Harm is thus reduced by a dignified and informed use of agency: those who wish to overcome the crippling effects of trauma know when and how to counteract it. Indeed, the *coping argument* appears highly favoured by defenders of trigger warnings in education (Manne, 2015; Gust, 2016; Dickman-Burnett & Geaman, 2019). Indeed, those who clearly deny promoting avoidance explicitly endorse trigger warnings as a way to cope. In the *New York Times*, Kate Manne (2015: 5) argues that:

---

[2] As one reviewer pointed out, this is not necessarily the case in the popular conversation. For example, one can find clear and explicit endorsements of the avoidance argument in online forums, such as Reddit.

> […] Increasingly, professors like me simply give students notice in their syllabus, or before certain readings assignments. The point is not to enable – let alone encourage – students to skip these readings or our subsequent class discussion (both of which are mandatory in my courses, absent a formal exemption). Rather, it is to allow those who are sensitive to these subjects to prepare themselves for reading about them, and better manage their reactions. The evidence suggests that at least some of the students in any given class of mine are likely to have suffered some sort of trauma, whether from sexual assault or another type of abuse or violence. So I think the benefits of trigger warnings can be significant.

In short, the warning enables students to "better manage their reactions" and therefore reduces psychological harm. Here, harm reduction focuses on those who have already suffered and are therefore at risk to be 'triggered'. Thus, to succeed is to make the worse-off better off.

Again, coping techniques differ from safe spaces. The latter are reactive. They help those who are unable or unwilling to further engage with certain materials. It follows that safe spaces are an *exit* from academic engagement rather than an aid to it.[3]

However, reducing harm is not the only objective. Psychological distress also represents an obstacle to learning. We can build on the coping argument to formulate the *learning argument*. Warnings not only help reduce suffering but also improve learning. Consider two passages. "When students request trigger-warning accommodations, they are informing educators about the importance of the nature of their experience and what they need to fully engage in an academic space" (Byron, 2017: 118). It is hard to interpret "fully engage in an academic space" independently of learning without losing sight of what academic institutions are. Similarly, we find an analogy with accommodations for disabilities.

> A trigger warning does not give permission for students to skip class, avoid a topic or choose alternative readings. What it does is signal to survivors of abuse and trauma that they need to keep breathing. It reminds them to be particularly aware of the skills and coping strategies that they have developed and to switch them on.

Triggers warnings are necessary adjustments for students who hold in their bodies one of the most prevalent but also most disabling of wounds – trauma. Like adjustments for dyslexia, they do not solve the challenge of being different, they simply make it easier to navigate the difficulty of living in a world that assumes certain norms (Gust, 2016).

Thus understood, trigger warnings do more than reduce suffering; they improve learning. Psychological distress negatively affects both our well-being and our ability to learn. Trigger warnings thus help the traumatised learn.

---

[3] One might distinguish between reactive and proactive safe spaces. A reactive one is the kind of room provided for those distressed by something and who have withdrawn from a conversation. A proactive safe space would require that one announce that one is in a safe space in order to engage in a difficult conversation.

The case for trigger warnings is primarily, if not exclusively, teleological. It is unclear if anyone considers a non-teleological argument to be decisive. We find claims about "more meaningful academic experience" or "respectful discourse", but such remarks are neither fully spelled out nor central to the argument. It is hard to find anyone arguing that *even if* trigger warnings did not substantially help, their non-teleological value would outweigh any stifled speech or the failure to prepare students for a harsh world.

3. Part III: The Evidence

In the last few years, empirical studies have emerged on the effects of trigger warnings. A study published in 2019 found that "[…] our meta-analytic confidence intervals were narrow, suggesting high precision – yet still showed trigger warnings plausibly have no effect or might even work slightly in the direction of causing harm" (Sanson et al., 2019: 790).

Another paper, published in 2020, specifically studied the impact of trigger warnings on their target population: *people who have been traumatised*. This design feature ensured that the "trauma naïve" were not muddying the waters by opining on experiences that are foreign to them. This study found that:

> Trigger warnings did not reduce anxiety for this sample broadly. Trigger warnings also did not reduce anxiety among people who met a clinical cutoff for PTSD symptoms, reported a diagnosis of PTSD, or reported that the stimuli matched the content of their past trauma. Trigger warnings showed trivially small effects on response anxiety overall. When effects did emerge, they tended towards small *increases* in anxiety rather than decreases (Jones et al., 2020: 914).

To avoid proceeding piecemeal or producing an unrepresentative sample, we turn to the first and only meta-analysis conducted on the effects of trigger warnings. Surveying twelve papers, the meta-analysis found no evidence that trigger warnings work as their advocates claim. They tested the effects of trigger warnings on four dimensions: 'response affect' or emotional reactions to the material, avoidance of the content, anticipatory anxiety, and comprehension of the material. A few numbers will provide a better understanding of the scope of the meta-analysis. Nine articles included studies testing for response affect, five articles included studies testing for avoidance, five articles included studies testing for anticipatory anxiety, and three articles contained studies testing for comprehension. We further note that four out of the twelve articles contained the findings of more than one study. Overall, the mean number of participants per study was 298.[4]

The results are unambiguous: not a single study found clear positive effects. Trigger warnings neither help individuals cope by regulating their emotions, nor do they result in more avoidance, nor do they help comprehension or learning. However, they found that:

> Overall, we found that trigger warnings had no meaningful effect on response affect, avoidance, or educational outcomes (i.e., comprehension). However, trigger warnings reliably increased anticipatory distress before viewing material (Bridgland et al., 2023: 16).

---

[4] A table of the studies contained in the twelve articles can be found at Bridgland et al. (2023: 757–760).

While it is clear that trigger warnings do not help, the extent to which they harm is more complicated. As mentioned, they do raise anticipatory distress. However, this is "a small increase in negative emotions" (Bridgland et al., 2023: 17). Though anticipatory anxiety is undesirable and is perversely inflicted on those who need help, we should not exaggerate its gravity. Raising someone's anxiety levels is unpleasant and objectionable, but it is not the same thing as a panic attack or other acutely painful psychological states.

Moreover, the effect of trigger warnings on learning is complicated. Overall, they do no good because the results wash out: those who have suffered trauma see the lowest results after reading a trigger warning while those who have not been victimised score the highest (Bruce & Roberts, 2020: 163). Thus, trigger warnings do not generally help learning but they do manage to mildly harm the traumatised. Again, this is reverse-prioritarianism, or the harming of the worst-off. We note that only three of the articles report findings on comprehension.

We identify a final unhappy finding. One paper has found that trigger warnings worsen the aesthetic experience of visual art.[5] Independent of political views, everyone's aesthetic appreciation was worsened and negative emotions were raised when trigger warnings were provided prior to viewing visual art (Jones et al., 2023). If replicated, these results reveal two further costs. One, they reduce well-being by lessening our appreciation of culture and human excellence. Two, by negatively affecting aesthetic appreciation and raising negative emotion, trigger warnings become a tool for indoctrination. If everyone's artistic experience is lessened by trigger warnings, then their use seems in tension with the goal of cultivating independent judgment. By selectively using trigger warnings, one can manipulate which works are well received and which are not.

At this stage, one might challenge the claim that the evidence clearly favours one conclusion. For instance, one might wonder if we are relying too heavily on the above meta-analysis at the expense of other forms of evidence. Furthermore, one might want to find fault with the meta-analysis itself. We reply that this kind of scepticism is unwarranted.

To begin, we would need to know what kind of evidence remains unexamined and how probative it is. Our reading of the various pieces written in defence of trigger warnings reveals few if any appeals to empirical studies. Those who claim to have evidence tend to rely on self-reports; they claim that they work, that they produce a certain effect, because that was their personal experience (Nelligan, 2022). This is unfortunate because the limitations of self-reports are well known. Let us restate two of their many known flaws. First, they are subject to socially desirable responding, or the tendency to provide flattering self-descriptions. Second, they are vulnerable to recall bias, or the systemic failures to recall events or recall them accurately (Althubaiti, 2016).

Therefore, the fundamental problem is that the evidence used by defenders of trigger warnings, like self-reports, is far weaker than the evidence cited above. Faced with this situation, there are only two options available: to adopt either a Cartesian or a Pyrrhonian standard for scepticism. On the former, we reject a belief the moment we can identify any reason to doubt it. However, adopting this standard would be self-defeating because the defenders of trigger warnings face many compelling reasons to doubt, as the above meta-analysis makes clear. More generally, there is no point engaging with empirical evidence or the sciences in general if one is going to adopt a Cartesian standard, because careful scientific studies rarely, if ever, yield findings that are quite literally indubitable.

---

[5] This series of studies included 213 participants and the effect size was moderate at 0.44.

If we endorse a Pyrrhonian standard, then the defenders of trigger warnings are no better off. To suspend belief, we need to formulate two opposing arguments of equal strength. However, there is no good reason to weigh equally a meta-analysis and self-reports. Therefore, the arguments are of very unequal strength and we have not met the necessary conditions to suspend belief.

In short, appealing to unexamined evidence is a red herring because it camouflages the very unequal value of the evidence at hand. If anecdotes and self-reports were as epistemically valuable as well-designed and expensive studies, then why bother with the latter?

Nor is it reasonable to point to some limitations of the meta-analysis to cast doubt upon its value. While it is true that twelve articles might seem low for a meta-analysis, this is deceptive. First, we should not presume that one article contains one study; a quarter of the articles contain the results of several studies. Second, not only is an analysis of twelve articles non-negligible, but this limitation is explained by the topic's recency. It is desirable to conduct more studies, but that fact does not commit us to dismissing a significant amount of scholarship until we reach what appears to be some arbitrary threshold for scientific respectability.[6]

Therefore, we maintain that the available empirical evidence on the effectiveness of trigger warnings is remarkably consistent. For their advocates, the facts yield little consolation. The evidence is strongest in support of the claim that they do no good. There is good but less compelling evidence that they do some harm.

4. Part III: Uninformed Arguments and Unintended Consequences

We now evaluate the arguments in light of the evidence. Both teleological arguments fail because they are unsound. If we should use trigger warnings to help students cope with distressing materials, the problem is a lack of evidence, let alone dispositive evidence. Even if the coping argument were valid, it would be unsound. Trigger warnings do not help students cope and this means that they cannot have any instrumental value because a key premise of the argument is false.

The argument is unsound rather than speculative. Before such studies were conducted and published, these arguments may have been speculative. When no one knows the facts, we rely on plausible assumptions. However, once we have knowledge on the topic, we cannot responsibly ignore it. Converging studies by different research teams are a strong basis for knowledge. If one's claims are contradicted by such findings, then one is no longer speculating. There is a big difference between the claim "no one knows, this seems plausible" and "several studies have tried to establish this claim and all have failed". As the meta-analysis above clearly states, trigger warnings do not produce the therapeutic effects their proponents claim. As arguments that rely on one or more false premises are unsound, it follows that the coping argument is unsound.

Similarly, the learning argument claims trigger warnings improve learning outcomes; they help students cope and this allows them to "fully engage". Once again, the meta-analysis is clear: trigger warnings do not help the vulnerable learn. Again, assuming that the argument is valid, it remains the case that a key premise is undermined by the evidence. The learning argument is also unsound. Not a single study assessed in the meta-analysis found clear evidence for either effective coping to reduce emotional distress or effective coping to improve learning.

---

[6] We note that a study published after the meta-analysis finds similar results, see Kimble et al. (2024).

Finally, if we turn to the avoidance argument in our search for a harm reduction argument, we find the exact same problem. The evidence is again unfavourable. Trigger warnings are ineffective tools if one seeks to promote avoidance of distressing materials. In fact, not only do they fail to promote avoidance, but they do the opposite: they have a forbidden fruit effect and drive more engagement with the material (Bridgland et al., 2023: 16).

All the teleological arguments are felled by their reliance on false premises. Advocates of trigger warnings have not only failed to motivate key premises, but have failed to notice that the evidence is dispositive. Like trickle-down economics and horoscopes, there is no evidence that trigger warnings work.

Even if the argument fared better, we would still need an all-things-considered argument. As the studies make clear, trigger warnings can impact various outcomes. It is possible that trigger warnings produce both desirable and undesirable consequences. Suppose that trigger warnings raised anticipatory anxiety, but also improved learning. It would seem that we would need to investigate how much they raise anticipatory anxiety and how much they improve comprehension, and then we would need to determine the relative value and disvalue of these outcomes. So far, the defenders of trigger warnings have not only produced unsound teleological arguments, but they have argued as if all good things coincide. Empirically-informed normative arguments should rely on a less naïve worldview.

Ironically, there is compelling evidence that trigger warnings are harmful. As stated, twelve studies, appearing in five articles, found that they raise the anticipatory anxiety of those who have been traumatised. Moreover, there is some evidence that they worsen the learning outcomes for the traumatised. While trigger warnings have no discernible instrumental value, they do have measurable instrumental disvalue. Judged by their effects, trigger warnings are morally pernicious; they promote neither psychological well-being nor learning, and yet they manage to harm the only group they purport to help.

While the debate has focused on teleological arguments, we wish to briefly sketch out a non-teleological argument or two. We want to show that modelling the defence of trigger warnings on other established practices will provide neither an easy nor a quick fix.

First, we could imagine someone defending trigger warnings as a kind of anti-discrimination measure. Trigger warnings, like access ramps, ensure equal access. However, such an argument would fail because there is no sense in which the absence of trigger warnings makes anyone worse off. Because trigger warnings are ineffective, they equalise nothing, they offset no disadvantage. Thus, there is a disanalogy between trigger warnings and accommodations or anti-discrimination measures like access ramps. The latter are only valued because they do help level the playing field and enable the vulnerable to engage more equally. If anything, insisting on using an ineffective tool that harms the worse-off fails to include or treat them with their due respect.

Second, we could imagine someone defending trigger warnings as a form of recognition. We might model them on land acknowledgements, which are primarily expressive and symbolic gestures. While they might not make the traumatised better off in terms of avoidance, coping, or learning, we might insist that goods like recognition provide sufficient reason to use them. The problem is that this argument is essentially a *non sequitur*. Social recognition is valuable, but it is unclear why it should be achieved or is best achieved in the classroom. Time, place, and manner, as lawyers are fond of saying, need to be considered. And even if one were to argue that educational institutions should help us make the suffering of the traumatised more salient, this again misses the mark because trigger warnings are not essential to any public conversation on trauma or mental health. For instance, post-secondary students, particularly female ones, might worry about sexual assault and believe that it is cruelly downplayed in public conversation. They might find this oppressive. Still, this in no way explains why trigger warnings are the effective or proper way of combatting this social ill.

To recapitulate, the evidence we have examined puts the advocate of trigger warnings on the backfoot. For teleological arguments, key premises are contradicted by the evidence. Not a single research team can point to clear benefits in terms of response affect, avoidance, or learning outcomes. Regrettably, their most reliable effect is to raise the anticipatory anxiety of the traumatised. Our brief discussion of non-teleological arguments reveals that abandoning a justification that relies on the effectiveness of trigger warnings is not so easily done. Nor can we easily model a defence of trigger warnings on the use of access ramps or land acknowledgments.

Finally, because we critically assessed both teleological and non-teleological arguments, we can pin down our opponents. Some defenders of trigger warnings seem keen to shift the goalposts.[7] The debate focuses on one argument, it is rebutted, then the defenders of trigger warnings deny that their support ever hinged on that argument or claim that this failure is inconsequential because trigger warnings can fulfil some other function, and so on. Instead of proceeding piecemeal, we can simply identify the defence being offered. Given that we know that both roads lead to a dead end, we should be unbothered. Shift your view or change the focus, we know that both teleological and non-teleological arguments fail.

5. Part IV: The Enemies of Prudence

Need we add anything? The arguments rested on false premises; they are unsound. Advocates believed or hoped that certain facts obtained, but they were mistaken. Perhaps the takeaway is that good intentions alone are not enough. If the argument for adopting trigger warnings is to improve the world by bettering the classroom in various ways, then the evidence should lead us to firmly reject them, as they act as a kind of reverse-prioritarianism.

However, this story does not capture the distinct failures that led to rapid adoption and widespread use of an ineffective and possibly damaging pedagogical practice. In fact, this case study is a textbook illustration of how a key intellectual virtue, prudence, is undermined by failures of character and failures in institutional incentive structures. Reflecting upon these failures allows us to better appreciate the enduring value of the virtues in professional activity.

---

[7] We thank a reviewer for raising this worry. It is true that many public debates proceed in this unhappy manner. One refutes an argument only to be confronted by another as if the first part of the debate was irrelevant. When this happens, it appears that one's interlocutor is not committed to any serious inquiry. Rather it appears to be evidence that one is faced with someone dogmatically committed to a conclusion—*this must be so!*

Prudence is more than lofty goals or pure intentions. To be prudent requires being clever—that is to say, capable of habitually achieving one's aims through successful means-ends reasoning (Hursthouse, 2023: 45, 55–56). Thus, a prudent pedagogue is someone who habitually succeeds in achieving their benevolent aims. And if the issue facing a prudent pedagogue is the impact of trauma or mental health on the classroom, then successful action requires understanding both the problem and the appropriate solution. Because prudence is about attaining "practical truth", a prudent person must know the relevant facts (Hursthouse, 2023: 45). Yet, the previous sections have clearly shown that the advocates and users of trigger warnings know little about trauma and mental health, and they know little or nothing about the ineffectiveness of trigger warnings to help overcome these obstacles.

One cannot select the appropriate treatment without an accurate diagnosis. Yet, those who have written in favour of trigger warnings clearly lack an accurate diagnosis. In her *New York Times* piece quoted above, Kate Manne (2015: 5) explicitly compares trauma to phobia. She writes:

> Mr. Lukianoff and Professor Haidt also argue in their article that we shouldn't give trigger warnings, based on the efficacy of exposure therapy - where you are gradually exposed to the object of a phobia, under the guidance of a trained psychotherapist. But the analogy works poorly. Exposing students to triggering material without warning seems more akin to occasionally throwing a spider at an arachnophobe.

This passage is rife with error. First, phobias are automatically activated in the presence of fear-relevant stimuli (Mineka & Ohman, 2002: 932). Arachnophobes respond to highly specific and predictable fear-relevant stimuli, namely spiders or things that resemble them. Trauma triggers, however, need not be meaningfully connected to one's past trauma (Ehlers et al., 2004: 407). Someone traumatised by combat or by sexual assault need not be triggered by depictions of war or sexual assault. Many triggers are *temporally associated* with the event—the trigger merely happened to be contemporaneous. The gap between the traumatising event and the trigger can be astounding. Indeed, something like "a pattern of light, or tone of voice" (Ehlers et al., 2004: 407) will act as a trigger despite not being particularly connected to the traumatic event, whether it was a traffic accident, battle, or sexual assault. This explains why the traumatised experience triggers as sudden and unexpected; because there is no obvious connection. Finally, the lack of inherent connection between the trigger and the traumatic event makes the preparation of a list of trauma-triggers long and iterative.[8]

Second, the point of the comparison was to cast the refusal to use trigger warnings as unnecessarily cruel and like wilfully provoking someone's phobic response. Yet, given that anything can be a trigger, to avoid cruelty one would need to warn about anything from turning on the classroom lights, to making a sudden gesture, to using a certain tone of voice, and so on. But this raises a further issue—if anything can be a trigger, then nothing guarantees that the warning itself could not be triggering. While it is easy to avoid presenting the specific fear-relevant stimulus to a phobic person, it is almost impossible to avoid triggers for the traumatised. To assume that a trigger must be related to the trauma in the way that fear-relevant stimuli is related to the nature of the phobia betrays grave ignorance.

---

[8] Correspondence with Dr. Payton J. Jones.

Even if advocates of trigger warnings argued that this practice was experimental, they would still face criticism. Prudent practitioners of any profession should know that our knowledge and best practices need updating. That is to say, at some point the experiment comes to an end and we need to judge its success.[9] More generally, what we learned in training, recently or long ago, might be incomplete or outdated. If the prudent are to succeed, then they cannot regularly act on false or outdated information. The prudent must occasionally check and update their beliefs. However, it is difficult to find any advocates of trigger warnings who have revised their views as study after study invalidated key claims. The meta-analysis draws on papers published as early as 2018. Today, there is little evidence that authors cited in part II revised their views. The objection can be thus formulated: If the purpose of prudence is to successfully guide choice and action towards the human good, then no one can be prudent if they are culpably ignorant of relevant information to the achievement of the good pursued. Those advocating for or using trigger warnings are acting neither on the best available evidence nor any serious evidence. Tragically, it appears that they never sought these facts; the practice spread long before we had any reliable data.

One might find this too uncharitable. First, the practice was not invented on the spot. It had existed for years in online forums for those discussing trauma. Second, we cannot constantly update all of our beliefs. The task would be all consuming. All of us, even models of prudence, are acting on some incomplete and outdated beliefs. Third, it is worth considering that trigger warnings seem intuitive.[10] We warn our friends and colleagues when they might encounter potentially unpleasant company. We also urge others to brace themselves for something difficult.

Our response allows us to further discuss the failures involved in the adoption and continued use of trigger warnings. First, it is true that trigger warnings were not clandestinely invented by malevolent pedagogues. However, the fact that something is an established practice, somewhere, is poor evidence of its effectiveness. If trigger warnings are alleged to have a therapeutic aim, then we should only use them if we have compelling evidence that they work. Thus far, the evidence is overwhelmingly dispositive.

Naturally, pedagogues talk to each other about their experiences. Encouragement from peers and appreciative words from students have certainly led some educators to sincerely believe that trigger warnings work. However, we would suggest that pedagogues must also realise the weakness of this case. Anyone responsible for teaching critical thinking should realise the limits of anecdotal evidence and the risk of confirmation bias.

And while we readily admit that trigger warnings seem intuitive, it is hard to accept that this is how we treat serious mental health concerns. We do not see why common sense or intuition should provide acceptable guidance if we are dealing with something as serious and complex as trauma. Consider how the recent COVID-19 pandemic produced an unfortunate number of amateur claims about how to avoid infection or treat it. We see little reason to draw a principled distinction between those who improvised solutions to a serious respiratory illness and those improvising remedies for serious mental health issues. If trauma is worth taking seriously, if it warrants the attention of pedagogues and changes to their professional conduct, then it is far too serious to rely on the unverified self-reports of untutored amateurs. If mental health is a serious issue, it deserves serious evidential standards. A prudent person does not treat human well-being so cavalierly.

---

[9] This is illustrated by the reactions to the film *Blink Twice*, see Pearson (2024). The film included a trigger warning and yet was roundly criticised. In other words, one demands the use of trigger warnings, they are used, they fail, and then one declares that these were not good or real trigger warnings. This is the same reasoning pattern behind “real communism”. Communism is demanded, it is tried, it fails, and then its defenders claim that the failure was due to the fact that this was not good or real communism.

[10] We thank a reviewer for urging us to consider this point.

The second objection makes an important point. It is true that we cannot constantly update all of our beliefs. And it is true that we all are likely to act on some incomplete and outdated information. Overwhelmingly, we rely on indirect knowledge, and updating all our beliefs would be an interminable task.

However, here too there is a division of labour. No one expects pedagogues to constantly revise all of their beliefs. Rather, there is a legitimate expectation that pedagogues should be better informed on issues relating to their profession, just as we expect military officers to follow developments relating to war. A prudent pedagogue therefore needs to keep up to date with information relevant to their actions, namely teaching. If they have identified certain conditions, diseases, or handicaps that pose a challenge, then they should be trying to update their beliefs to properly respond to these challenges.

Nor is anyone asking pedagogues to be omniscient. But if pedagogues are going to promote and use new tools and claim that the justification is the well-being of students, then there is nothing unreasonable about expecting them to proceed on the basis of up-to-date information. Moreover, there is not an infinite number of controversial new pedagogical tools. Updating one's beliefs on one of the controversies *du jour* within one's profession is reasonable.

A further question is raised by the above argument. Given that trigger warnings have been used and promoted for about two decades and that serious empirical research only began emerging in 2018, what explains such widespread imprudence? Why would so many pedagogues endorse them so swiftly and uncritically? The best explanation, we fear, is the failure of intellectual and moral virtues aided and encouraged by institutional pressures.

We begin with conformism. There are two kinds of conformism, strategic and non-strategic (Hill & Garner, 2021). The latter is a decent kind of conformism. All of us must, more or less frequently, defer to the superior knowledge or judgment of others. When faced with a contagious virus, we should trust epidemiologists and virologists. At other times, it is wise to put our faith in numbers. The best reason to think that something is vulgar or boring is that so many believe so. Non-strategic conformism is the kind of pro-social deference and trust that all communities require to function well and that all of us practice to varying degrees.

If non-strategic conformism was at play, then pedagogues may have acted in good faith and on the best of intentions, but this is still a failure of intellectual virtue. Consider the non-strategic conformism that occurred during the COVID-19 pandemic. We deferred to experts of the relevant type. Evidence was presented and discussed. Yet, which eminent experts of trauma took up the cause of trigger warnings? Which studies were produced to support their claims? Deferring to others is only virtuous if one has good reason to trust their judgment. Given that pedagogues presumably have received university training, know how to search for evidence, and read scholarly papers, it is hard to imagine how such deference to either fashion or superiors lacking the relevant expertise amounted to a display of intellectual virtue. Teachers and professors are precisely those we trust to teach research skills and critical thinking. Yet, faced with very strong claims about crucial issues like mental health and learning, those presumably responsible for training others to think critically and to differentiate between expert and amateur opinion conformed with little hesitation. Their conformism might have been non-strategic, but it was also naïve and unjustified.

Strategic conformism is not motivated by genuine deference or trust. Rather, it is the conformism driven by self-interest. To conform for strategic reasons is to go along to get along, rather than to act on any ethical or epistemic principle. More precisely, strategic conformism is either motivated by a defensive aim to preserve one's position or an offensive aim to improve it. In both cases, the need to communicate that one conforms to the group's aim or ethos or ideology is at the heart of much virtue signalling (Hill & Garner, 2021). One says and does as others do because failure to conform poses a threat to one's reputation or status.

This brings us to a final point about the link between moral and intellectual virtues. Prudence without courage is of limited value. This is not to endorse the stronger thesis that is the unity of the virtues, that one must possess all of the virtues in order to possess any individual virtue. Rather, the claim is that one might genuinely possess a virtue, but if the lack of another inhibits the full exercise of that virtue, the value of the inhibited virtue is diminished. Consider prudence in the absence of courage. If some pedagogues came to the conclusion that trigger warnings were useless or unlikely to help, but conformed out of fear of social repercussions, then their prudence was inhibited by their unwillingness or inability to face risk. While some risks might warrant such behaviour, we take it that pedagogues should be made of sterner stuff and be able to face mild threats, such as public criticism or mockery. After all, if prudence is about achieving the human good, in this case helping us achieve the good for those who struggle with mental health, then it seems that fear for one's reputation or status should be overcome. And the relevant virtue to overcome fear and danger is courage.

Put otherwise, if strategic conformism is responsible for the adoption and dissemination of trigger warnings, then this means that a lack of courage is the deeper explanation.[11] Make no mistake, the case of trigger warnings presents us with trained professionals unwilling to face mild risks when it is their role to openly and honestly reason and teach others how to do so. To reprise the Aristotelian scale, the most charitable interpretation is that this repeated failure to face these mild risks is incontinence and the least charitable view is that it is vicious. Either way, the best explanation for the misuse of an intellectual virtue (prudence) is a failure of character (courage).[12]

---

[11] This presumes that strategic conformism is at play. If non-strategic conformism is at play, then imprudence rather than cowardice is the issue. However, we have already examined non-strategic conformism.

[12] We agree with a reviewer that we should not be unduly harsh. Undeniably, there are brave academics who stand up to wrongdoing by peers and superiors. However, it seems that the use of trigger warnings spread so quickly precisely because little resistance was offered. If we must appeal to what working academics witness, as the reviewer rightly urges us to do, then we must also mention a lack of courage. "I agree with you, but cannot say so publicly" has become sadly banal in academia.

None of the above is a denial of the social structures, and more precisely the institutions, that have played a role in these failures of virtue. Excellences of character or of the mind are not cultivated or corrupted in a vacuum. Schools and universities share in these failures of character and of the mind. A useful way to explain their corrupting effects is to draw on Alasdair MacIntyre's (1981) famous analysis of virtues, practices, and institutions. Virtues are the forms of excellence necessary to achieve the goods that are internal to a social practice, a cooperative social activity with internal norms of excellence. To excel at chess or farming, to flourish as a chess player or a farmer, one needs to develop certain virtues. However, complex social practices occur within a broader social world and the structures that sustain them have their own goods that are external to the practice, like fame, money, or power. One might want to excel at chess and one may also need to interact with chess clubs, federations, and so on. Because these are institutions, they have aims other than the excellence of the practice, such as being profitable or famous. Thus, there is a constant tension between the goods internal to the practice and the goods external to it but with which relevant institutions are concerned.

Schools and universities are institutions. Whether the specific practice is scholarship, teaching, or something else, this is a distinction that makes little difference. If trigger warnings, or any other ill-justified pedagogical tool, are promoted by one's hierarchical superiors, if these are presented as part of a broader ideology endorsed by those who hire, promote, and fire, then individual members of the profession may need to choose between the goods internal to their practice, such as intellectual rigour or integrity, and the external goods, such as their salaries. As many in education already know, this picture, whilst sadly true, is incomplete. Teachers and professors are not simply confronted with their immediate superiors. External institutions, such as ministries of education or various institutions distributing research funds, can and do exert pressure to conform with particular ideologies or at the very least signal that one acquiesces. The descriptions of funds and the application forms make this plain.

Additionally, students and parents are increasingly treating education as a transactional relation in which payment rather than expertise determines which choices should be made. This means that pressure to conform is not only coming from above, by department chairs, deans, or national funding agencies, but also from below, as students and parents consider themselves qualified to judge and criticise the professional conduct of pedagogues. Unfortunately, these two forms of pressure can be mutually reinforcing.

In other words, the institutional design of schools and universities does not favour the cultivation of virtue. Instead, the incentive structure and the pervading ideology encourages conformism, strategic or non-strategic. The spread of trigger warnings, ineffective and harmful, is the result of a failure of prudence on the part of pedagogues. However, the failure of pedagogues to be prudent or to voice their prudent thinking must be explained by other failures. In part, it is a failure of courage. Members of a profession who are fond of slogans like "Speaking truth to power" have often failed to follow through. In part, it is the predictable result of poor institutional design and culture. For all their rhetoric about independence of thought or speaking truth to power, or the need to challenge everything, schools and universities are institutions moved by external goods.

A final objection still stands. One can condemn trigger warnings and also reject our characterisation of the failure. This error does not prove that pedagogues were imprudent. After all, Aristotle (2009: 1109b15–20) argued that small deviations from goodness were not to be blamed, but only wide ones. Trigger warnings are but one pedagogical tool and we ought not draw overly strong conclusions on the basis of the failures of those who have promoted or used them.

The problem is that trigger warnings do not appear to be an outlier. We can find multiple examples of ill-founded beliefs and practices in contemporary pedagogy. In each case, we find that the beliefs or practice spread despite the fact that there was either little or no reason to endorse them. Like trigger warnings, some of these beliefs or practices seem costly or mildly harmful to students. And like trigger warnings, these ill-founded beliefs and practices were adopted, with little resistance, by educated pedagogues who should have known better than to uncritically endorse them.

Learning styles, the notion that there is no single most effective way to learn, but that different students have different learning styles and that teaching should reflect a diversity of approaches to learning, had been promoted for decades before it was adequately tested and the results made clear. And the results are very clear: while people may prefer to learn in a certain manner, there is no evidence that they learn any better with their preferred method. In short, there is no good evidence for learning styles (Pashler et al., 2008; Cuevas, 2015).

We also witnessed strong disagreement over how to teach students to read, the so-called “reading wars” (Sohn, 2020). Here too, the empirical evidence did not and does not stack equally. The old-fashioned and maligned phonics approach has been vindicated by numerous studies and now a meta-analysis (Jeynes, 2007; Report of the National Reading Panel, 2000; Report of the National Early Literacy Panel, 2008). There might have been conjecture for both sides, but there never was strong evidence that the two rival approaches were equally effective.

Even more recently, we have seen a drive to include screens and ‘modernise’ teaching. And yet again, we see that university-educated pedagogues either sincerely endorsed or silently acquiesced to unproven methods and weak arguments. Today, countries that led this trend, like Sweden, are backtracking in the face of damning evidence (Associated Press, 2023; Deconinck, 2023).

In short, the failures behind trigger warnings were not exceptional. It would appear that, time and again, educated professionals have failed to critically assess beliefs and practices relevant to their craft. And at the risk of repetition, these beliefs and practices were promoted and endorsed despite a lack of empirical evidence. What we are offering is essentially an inference to the best explanation. Pedagogues were not misled by fraudulent publications or poorly designed studies. We believe that failures of prudence and failures of courage, as well as institutional pressures to conform, provide a very compelling explanation as to why primary and secondary school teachers and college and university professors have so often failed to live up to the standards they profess. Aristotle would be unsurprised. Virtue, moral and intellectual, was always a question of both individual achievement and good institutions.

Author Contributions: R.-P.G. (90%) led the analysis and writing of the paper with contributions from N.E.J. (10%).

Funding: This research was funded by the University of British Columbia, Grant ID GR024028.

Acknowledgments: We wish to thank, in no particular order, Jeremy Reid, Andrew Irvine, Christopher Martin and Julien Picault.

Conflicts of Interest: There are no conflicts of interest for either author.

References

Althubaiti, A. (2016). Information bias in healthcare research: Definitions, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare*, *9*, 211–217. [CrossRef] [PubMed]

Aristotle. (2009). *Nicomachean Ethics*. Oxford University Press.

Associated Press. (2023, September 11). Switching off: Sweden says back-to-basics schooling works on paper. *The Guardian*.

Bridgland, V. M. E., Jones, P. J., & Bellet, B. W. (2023). *A meta-analysis of the efficiency of trigger warnings, content warnings, and content notes*. OSF Preprint. [CrossRef]

Bruce, M., & Roberts, D. (2020). Trigger warnings for abuse impact reading comprehension in students with histories of abuse. *College Student Journal*, *54*(2), 157–168.

Byron, K. (2017). From infantilizing to world making: Safe spaces and trigger warnings on campus. *Family Relations*, *66*(1), 116–125. [CrossRef]

Cuevas, J. (2015). Is learning styles-based instruction effective? A comprehensive analysis of recent research on learning styles. *Theory and Research in Education*, *13*(3), 308–333. [CrossRef]

Deconinck, C. (2023, August 22). Sweden to cut use of screens in schools as reading standards suffer. *Brussels Signal*..

Dickman-Burnett, V. L., & Geaman, M. (2019). Untangling the trigger-warning debate: Curating a complete toolkit for compassionate praxis in the classroom. *Journal of Thought*, *53*(¾), 35–52.

Ehlers, A., Hackmann, A., & Michael, T. (2004). Intrusive re-experiencing in post-traumatic stress disorder: Phenomenology, theory, and therapy. *Memory*, *12*(4), 403–415. [CrossRef] [PubMed]

Gust, O. (2016, June 14). I use trigger warnings - but I'm not mollycoddling my students. *The Guardian*.

Hill, S., & Garner, R.-P. (2021). Virtue signalling and the condorcet jury theorem. *Springer Link*, *199*, 14821–1484. [CrossRef] [PubMed]

Hursthouse, R. (2023). Practical wisdom: A mundane account. In *Virtue and action: Selected papers* (pp. 45–65). Oxford University Press.

Jeynes, W. (2007). A meta-analysis of the relationship between phonics instruction and minority elementary school students academic achievement. *Education and Urban Society*, *40*(2), 151–166. [CrossRef]

Jones, P. J., Bellet, B. W., & McNally, R. J. (2020). Helping or harming? The effects of trigger warnings on individuals with trauma histories. *Clinical Psychological Science*, *8*(5), 905–917. [CrossRef]

Jones, P. J., Bridgland, V. M. E., & Bellet, B. W. (2023, May 25). *Content warnings reduce aesthetic appreciation of visual art*. Psychology of Aesthetics, Creativity, and the Arts.

Kilpatrick, D. G., Resnick, H. S., Milanak, M. E., Miller, M. W., Keyes, K. M., & Friedman, M. J. (2013). National estimate of exposure to traumatic EVENTS AND PTSD prevalence using DSM-IV and DSM-5 criteria. *Journal of Traumatic Stress*, *26*(5), 537–547. [CrossRef] [PubMed]

Kimble, M., Koide, J., & Flack, W. F. (2024). Students responses to differing trigger warnings: A replication and extension. *Journal of American College Health*, *72*(6), 1668–1671. [CrossRef] [PubMed]

Lukianoff, G., & Haidt, J. (2015). The coddling of the American mind. *The Atlantic*, *316*(2), 42–52.

MacIntyre, A. (1981). *After virtue*. Notre Dame University Press.

Manne, K. (2015, September 19). Why I use trigger warnings. *The New York Times*.

Mannix, L. (2022, November 29). Trigger warnings are everywhere, but do they work? *The Sydney Morning Herald*.

Mineka, S., & Ohman, A. (2002). Phobias and preparedness: The selective, automatic, and encapsulated nature of fear. *Biological Psychiatry*, *52*(10), 927–937. [CrossRef] [PubMed]

Nelligan, M. (2022). My case for using and respecting trigger warnings. *National Alliance on Mental Illness*. consulted January 2024. link to the article.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, *9*(3), 105–119. [CrossRef] [PubMed]

Pearson, V. (2024, September 9). 'Blink twice' prompts discourse about censorship and trigger warnings. *The Hilltop*. link to the article.

Report of the National Early Literacy Panel. (2008). *Developing early literacy: A scientific synthesis of early literacy development and implications for intervention*. link to the article.

Report of the National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. link to the article.

Rubin, D. (2008). A memory-based model of posttraumatic stress disorder: evaluating basic assumptions underlying the PTSD diagnosis. *Psychological Review*, *115*(4), 985–1011. [CrossRef] [PubMed]

Sanson, M., Strange, D., & Garry, M. (2019). Trigger warnings are trivially helpful at reducing negative affect, intrusive thoughts, and avoidance. *Clinical Psychological Science*, *7*(4), 778–793. [CrossRef]

Schein, J., Houle, C., Urganus, A., Cloutier, M., Patterson-Lomba, O., Wang, Y., King, S., Levinson, W., Guérin, A., Lefebvre, P., & Davis, L. L. (2021). Prevalence of post-traumatic stress disorder in the United States: A systematic literature review. *Current Medical Research and Opinion*, *37*(12), 2151–2161. [CrossRef] [PubMed]

Schnurr, P. P. (2025). *Epidemiology and impact of PTSD*. National Center for PTSD. link to the article.

Shulevitz, J. (2015, March 21). In college and hiding from scary ideas. *The New York Times*.

Sohn, E. (2020, April 26). It’s time to stop debating how to teach kids to read and follow the evidence. *Science News*.