

Tortured Logic: Information and brutality in interrogations

John W. Schiemann PhD*

Key points of interest:

- Employs mathematical game theory to assess the pragmatic argument for interrogational torture, with the intuition behind the math presented here.
- Multiple real world uncertainties facing both interrogators and detainees mean that information from torture will be unreliable.
- These same uncertainties cause torture to exceed limits and controls imposed by torturers themselves, leading to slippery slopes of torture frequency and brutality.
- A rigorous logical examination supports historical, social scientific, psychological, and neurobiological evidence refuting the pragmatic argument for interrogational torture.

Abstract

Background: Pragmatic arguments for interrogational torture rest on the twin assumptions that torture generates reliable information and that torture can be controlled and limited. **Methods:** I

assess the claims of torture proponents by providing the intuition behind a game theoretic model of interrogational torture. Tracing out the logic of different combinations of possible interrogators and detainee types results in eight outcomes that can be compared to three claims made by torture proponents: that information will be predictably reliable, that the frequency of torture will be minimized, including no torture of innocents, and that the severity of torture can likewise be limited and controlled. **Findings:** Of the eight outcomes generated by the model, only two result in full information, but an innocent is tortured in both and in one the detainee providing information is tortured after having no more information to give. Moreover, these outcomes are only possible for an extremely restricted and empirically unlikely combination of circumstances. With respect to torture frequency, detainees are tortured in seven of the eight outcomes, including innocent detainees. The incentives facing interrogators also compel them to ratchet up their brutality in an effort to compel information. **Discussion:** The outcomes of a model of interrogational torture based on the proponent ideal violate the three conditions individually necessary to support that ideal: (1) information from torture is unpredictable and unreliable, with no information and false information far more

*) Department of Social Sciences & History, Fairleigh Dickinson University, Madison, New Jersey.
Correspondence to: jws@fdu.edu

likely than good information; (2) torture will be used more frequently—including against innocents—than control and limits permit; (3) torture will be more brutal than controls and limits allow. **Conclusion:** The only thing reliably effective about interrogational torture is its ability to generate slippery slopes of frequency and brutality, violating the basic premises of the pragmatic argument for interrogational torture.

Keywords: interrogation, torture, game theory

Introduction

A recent international survey showed that over one in three respondents support torture to obtain intelligence in some circumstances (International Committee of the Red Cross (ICRC) 2016, 10). This is a decline from two-thirds saying no in 1999 to 48% in 2016 (ICRC 2016, 10). This change in public opinion is complemented and supported by an analogous shift to a pragmatic approach assessing costs and benefits in public policy discussions, legal thinking, and philosophy. Even the current President of the United States openly advocated interrogational torture “worse than waterboarding” partly because “it works” (Johnson 2016).

Although many understandably continue to attack interrogational torture on legal and rights-based grounds, others from philosophy, history, and public policy have challenged the pragmatic argument on its own terms (Johnson & Schmidt 2016, pp 132, 122; Rejali, 2007, esp. chaps. 21-22; Rumney, 2014). One of the most damning challenges to the proponent argument is one familiar to readers of this journal: the powerful neurological and physiological evidence that torture is not only unlikely to facilitate the recall of episodic memories, but may very well cause the destruction of such

memories via tissue loss in relevant brain regions (O’Mara, 2015).¹

I complement these arguments by approaching interrogational torture in a new way: examining the logic of interrogational torture using game theory, a branch of applied mathematics that models strategic interaction between two or more people.² In the limited space here I provide the intuition behind the formal argument.³ I argue that tracing out the proponents’ own logic of torture shows that torture is ineffective yet inevitably results in brutality that both outweighs any information gained and exceeds proponents’ own limits and controls. In other words, the outcomes of an analytical model of interrogational torture constructed according to the proponent’s own conception, with limits on torture, fail to match up to the claimed outcomes of the proponent normative model: what they say should happen when torture is introduced into interrogations.

The paper proceeds as follows. In section one I identify some necessary conditions

¹ See also O’Mara (2011, 2009), Elbert et al. (2011), Elbert et al. (2006), and Maercker & Forstmeier (2011).

² For an accessible introduction to game theory see Gibbons (1992). My analysis applies strictly to interrogational torture, where the goal is to compel information, but to ease the exposition I sometimes use “torture” alone. For another recent application of game theory to the effectiveness of interrogational torture, see Baliga and Ely (2016).

³ As a result, the presentation here is necessarily incomplete and imprecise in places. For the formal mathematical argument, including a description of the model and proofs of equilibria and propositions, see Schiemann (2016). I also address there in detail the understandable concern about using a rational model to account for the effectiveness aspect of interrogational torture, something I unfortunately lack the space to do here.

and the proponent's ideal outcome - four criteria of success - that interrogational torture is supposed to satisfy on the proponent account. This is the normative model. Having delineated in normative terms how interrogational torture should work on the proponent view, I first build an informal model of interrogational torture in section two. This is the analytical model. Informed by proponent ideas about limits and controls on torture, this model identifies the important elements in the strategic dynamic between a torturer and a detainee. The remainder of section two follows the proponent logic of torture to its conclusions, showing how different outcomes (e.g. a detainee provides information but is tortured anyway) emerge from the initial conditions. Section three confronts the normative model's pragmatic criteria of success sketched in section one with these outcomes of the analytical model. This comparison demonstrates that interrogational torture fails to satisfy the necessary conditions identified by proponents themselves.

The Proponents' Normative Model

Whatever the influence of utilitarian philosophers, it is legal scholars who have elaborated the most concrete proposals for institutionalizing torture in such a way that it maximizes information while simultaneously limiting abuses. They boil down to three individually necessary conditions an interrogational torture program must meet:

Condition 1. Information Reliability. Interrogational torture is successful if and only if detainees give up (nearly) all their information so that the ratio of clear and valuable information to all other information (non-valuable, false and misleading, and no information) is high.

Condition 2. Torture Frequency.

Interrogational torture is successful if and only if torture is not employed too frequently, that is, the total frequency of torture is low, torture is not used on cooperative detainees after they have provided all their information, and not on innocent detainees tortured for telling the truth.⁴

Condition 3. Torture Severity.

Interrogational torture succeeds only if torture is not employed too severely, that is, not much beyond the minimum degree necessary to compel valuable information.⁵

Put together, this is the ideal outcome claimed by proponents of interrogational torture:

Torture Justification Outcome. Torture in interrogations is justified on the proponent view if and only if torture is not used against cooperating detainees who have provided all their information, nor against innocent detainees, nor exceeds the minimum frequency and severity necessary, to elicit valuable information, and (the threat of) torture generates all, or nearly all, the valuable information possessed by knowledgeable detainees.

These are the necessary conditions to be satisfied by the proponents' normative model. We now sketch a game theoretic, analytical model reflecting the implicit logic of torture proponents.

⁴ Here and hereafter, all phrasing about "innocents," "necessary" and "unnecessary torture," and "minimum" torture is meant to capture the proponent point of view, not the author's.

⁵ Exactly what constitutes the "minimum" is unclear even for torture proponents, but a gross violation of any reasonable notion of minimum is sufficient to refute the pragmatic model.

Methods and Findings

Methods: An Analytical Model of Interrogational Torture

To think about torture from the proponent's perspective, imagine a detainee facing an interrogator willing to torture, but preferring not to do so.⁶ The detainee, of course, prefers not to be tortured, but some detainees (cooperatives) are willing to provide information if threatened with (more) torture whereas some detainees (resistants) continue to resist, refusing to provide information (or providing false information) even when threatened with torture and actually tortured. An innocent detainee resembles the cooperative type insofar as he is willing to lie and tell the interrogator what she wants to hear to in order avoid torture. Of course neither a cooperative nor an innocent detainee would want to cooperate with an interrogator and be tortured anyway.

Note that an innocent detainee can please the interrogator only if the interrogator asks a leading rather than objective question. An innocent detainee cannot provide an answer to the question "what is the address of the safe house?" but can say "yes" to the question "the safe house is at 10 Maple Lane, isn't it?" Of course, a detainee who actually knows something might confirm the leading question of 10 Maple Lane when he knows it is actually at 25 Elm Street. Thus the value of "cooperating" under leading questioning is entirely dependent upon the accuracy of the original question and so cannot—by definition—provide any new information. Despite numerous attempts throughout the

history of torture to prevent them, leading questions are used all the time and so we will include this possibility in our model.

Although proponents of interrogational torture always assume the detainee is a terrorist, rather than a suspected terrorist, the history of torture past and present is replete with examples of innocents swept up and tortured even when efforts are made to prevent this. Rejali's extensive analysis of torture by the French during the Battle of Algiers found that the most charitable ratio of the numbers of innocent who were arrested for every actual insurgent was a staggering 15 to one (Rejali 2007, p 483). Former U.S. Army general Janis Karpinski, commander of Abu Ghraib during the height of the abuses there, later estimated that "about 90 percent" of the detainees brought in for interrogation as terrorists were "innocent of terrorism or any related activity" (Karpinski 2005). Many of the detainees held by the United States at its military prison at Guantánamo Bay, Cuba were innocents sold to the US military for bounties in Afghanistan (Denbeaux et al., 2006, pp 15, 21). Thus, of the original 780 detainees, 730 have been transferred and 41 remain in custody because they are believed threats (New York Times, 2017). More than one in five of the known 119 detainees in the CIA's interrogational torture program were innocent (United States Senate 2014, pp 19, 42). Although in practice few interrogators are likely to place much weight on this probability, we want to assume our theoretical interrogators at least consider the possibility in order to give proponents the benefit of the doubt. Thus, we assume that if an interrogator really does think a detainee were innocent, she would refuse to torture.

What else might you be uncertain about as an interrogator? Suppose you have asked some objective questions and

⁶ To keep them distinct, I use the pronoun "he" for the detainee and "she" for the interrogator.

received good information in return. And then your detainee says he knows nothing more. Do you threaten (more) torture? In principle you would assume there must be some questions to which a detainee would have no answer; no detainee knows literally everything. If, however, you think he is holding back, you are committed to torturing more.

What will the detainee be uncertain about? Any cooperative or innocent detainee will have at least one major concern: that talking really will stop the torture (remember that the resistant type never talks or provides only false information). How likely is it that someone who has been kidnapped, thrown into a cell and tortured or threatened with torture will believe the interrogator that she will not torture if she is told what she wants to hear? It would be difficult for an interrogator to convince someone that she is not sadistic and only wants information. It is important to note that in order to give the proponents the benefit of the doubt, we will assume that the interrogator is *not* sadistic but instead pragmatic when we analyze the outcomes. We must, however, include the sadistic type of interrogator in the model because a real-life detainee would think it is at least possible that the interrogator is sadistic and not pragmatic and this could affect his behavior.

There is another problematic source of uncertainty that is not shared by both the detainee and the interrogator. While a cooperative detainee will assume that if he does provide good information it will be recognized as such—what is the point of torture otherwise?—this does not always happen. Sometimes interrogators do not understand that the information is valuable. It may contradict other (bad) intelligence in which they have confidence or it may seem implausible. Whatever the reason, the

interrogator may think the detainee is lying and so continue to torture—presumably to the surprise of the cooperative detainee.

Thus we have a detainee sitting across from an interrogator. The detainee could have information and be cooperative or resistant, or he could not have information and be innocent. The interrogator, who could be pragmatic or sadistic in the mind of the detainee, will have some estimation of the relative likelihood of each detainee type going in. Since, according to the logic of torture, it is always the threat of (more) torture in the future which is supposed to compel truth telling in the present, the detainee either provides valuable information or does not, and then the interrogator either tortures or does not. The detainee's decision to provide information depends on the value of the information (or the desire not to tell a lie if innocent), the anticipated costs of torture, and his belief about whether or not the interrogator is pragmatic or sadistic. Not providing valuable information might mean providing throw away, non-valuable information, providing misleading information, or no information at all—literally staying silent. If valuable information is provided, the interrogator might be uncertain about how valuable it is and will definitely be uncertain about whether the detainee has any more information he is not divulging. Using, in part, the detainee's behavior to update her beliefs about these uncertainties, the (pragmatic) interrogator will decide whether or not to torture. As spare as this sketch may seem, this model captures some of the most critical dimensions of the proponent model of interrogational torture.

Findings: Outcomes of the Analytical Model

To analyze interrogational torture, we systematically explore each combination of detainee and interrogator choices and beliefs

Table 1: *Interrogational Torture*

Objective Questioning	Leading Questioning
Complete valuable information, surprise torture	Ambiguous information, selective torture
Incomplete valuable information, surprise torture	False confirmation, selective torture
	Complete valuable information, selective torture
	Incomplete valuable information, selective torture
	No information, torture
	No information, no torture

and trace out what each type of interrogator and each type of detainee would do in that circumstance. Doing so results in eight outcomes, summarized in Table 1 in terms of both information and torture. Two occur under leading questioning only (*Ambiguous information, selective torture* and *False confirmation, selective torture*) and two occur under objective questioning only (*Incomplete valuable information, surprise torture* and *Complete valuable information, surprise torture*). The remaining four outcomes occur under both objective and leading questioning (*Incomplete valuable information, selective torture*; *Complete valuable information, selective torture*; *No information, torture*; *No information, no torture*).

To see how these outcomes are derived, suppose, for example, that a cooperative detainee faces a pragmatic interrogator and that the detainee is convinced that the interrogator really is pragmatic and so will not torture if he gives up information. The interrogator asks some objective questions and the detainee responds truthfully, providing all the valuable information he has because of the threat of (more) torture. The interrogator, for her part, is convinced that the detainee is cooperative, i.e. has information and gives it up under the threat of (more) torture.

The only remaining question is whether she believes that the detainee has given

her all the information he has or whether she believes he is still holding back. If she believes he has given up everything, then she does not torture (remember, she's pragmatic not sadistic). If she believes instead that he has *not* divulged all he knows and he is holding back information, she tortures him. In this case, the detainee will be surprised, since he believed the interrogator would believe he had divulged everything and so expected he would not be tortured. Note that this surprise is only from the perspective of the detainee and is based purely on what he thought the interrogator would think and do; it is not based on having necessarily actually given all his information. He could be surprised at the interrogator's response even though he was withholding information because he thought he had successfully deceived the interrogator. So there are really two outcomes here: one in which a detainee withholding some more information is unexpectedly tortured and one in which he really did give up all his information and is unexpectedly tortured, again with "unexpected" referring to the detainee's perspective: *Incomplete valuable information, surprise torture* and *Complete valuable information, surprise torture*, respectively.

The same two possibilities are also true when the detainee provides information and the interrogator does not torture because she really does believe that the

detainee has given up all his information. It might be the case that she is correct and there really was no withholding (*Complete valuable information, selective torture*) or it might be the case that she is wrong and the detainee got away with it, convincing the interrogator that he had given up everything he knew when in fact he had not (*Incomplete valuable information, selective torture*). In abbreviated form the four outcomes here are: (1) gave up everything, no torture, (2) withheld information, no torture, (3) gave up everything, torture, and (4) withheld information, torture.

It is worthwhile pointing out that only the first outcome has the potential to support the proponent point of view (more on why only “potentially” below). In the remaining outcomes, either information remains hidden or the detainee was tortured “unnecessarily” even on the proponent view. In order to get the four outcomes above, the detainee had to believe the interrogator was pragmatic, not sadistic and the interrogator had to believe that the detainee was cooperative and not innocent. The difference between the two selective torture and the two surprise torture outcomes was driven by whether the interrogator believed the detainee had given up everything (and so did not torture the cooperative detainee) or whether she believed the detainee was withholding information (and so tortured the cooperative detainee).

Thus far we have identified four outcomes of the model. Changing other beliefs results in other outcomes. Working through the different possible combinations of questioning type as well as beliefs about possible detainee innocence, interrogator type, and full disclosure results in the following possible additional outcomes: *Ambiguous information, selective torture; False*

confirmation, selective torture; No information, torture; and No information, no torture.

The first two of these additional outcomes occur under leading questioning only. In both outcomes the interrogator has asked a leading question and received the answer she wanted to hear. To return to our earlier example, we imagine she asked “the safe house is at 10 Maple Lane, isn’t it?” and the detainee said “Yes.” In the Ambiguous information outcome the interrogator is unsure whether the detainee is truly innocent or is cooperative and really does have information because all she heard was “Yes” to the leading question. Either detainee type might have replied “yes.” Either way, she is satisfied because she got the answer she wanted to hear from her leading question and so does not torture (though she would torture if displeased with the answer, hence the “selective torture”). The “ambiguous” in the name of this outcome captures the uncertainty about the nature of the detainee (i.e. innocent or cooperative). In the False Confirmation outcome, the interrogator actually knows the detainee is truly innocent, but just does not care because she is again happy with the false confirmation or confession she received and so does not torture him. If, however, the detainee refused to say what the interrogator wanted to hear, she would torture him. This outcome, in other words, corresponds to the vast majority of torture cases, cases in which suspected criminals are tortured for confessions rather than terrorists tortured for intelligence.

The two no information outcomes occur under both leading and objective questioning. In the *No information, torture* outcome, neither an innocent nor a cooperative detainee believes the interrogator’s promise not to torture if the detainee cooperates and so both

types refuse to cooperate (the resistant type never provides information).⁷ For her part, the interrogator believes the likelihood the detainee has information (i.e. is not innocent) is high enough for her to torture in an attempt to compel the information from what she believes is a detainee attempting to hide information. In the *No information, no torture* outcome, the interrogator thinks the chance that the detainee is innocent is high enough that she decides not to torture, despite not receiving any valuable information.

Discussion

Before turning to how these outcomes compare to the necessary conditions of the normative model in section one above, it is important to note two caveats and limitations that come immediately to mind. First, I do not claim that the model sketched here captures everything about torture. Many, many aspects of torture are better captured—to the extent they can be captured at all—by victim testimonies like Gestapo torture and Holocaust survivor Jean Améry, by paintings such as those by Khmer Rouge torture survivor Vann Nath, or by Wisława Szymborska's "Tortures" and other poetry, just to name a few (Améry, 1980; Szymborska 1998, pp 202-203). Second, I do not even claim that the model captures everything about the effectiveness of interrogational torture for information. For example, while some victims may provide information as the result of a rational decision and expectation torture will stop, I

recognize that others may do so from a loss of control and what might be a baseless hope the torture will stop. Even here, however, the model can serve an "as-if" function, capturing victim and torturer behavior in the same way that physics models billiard balls even if players fail to make the calculations in their head. Either interpretation permits us to interrogate the logic of those who argue torture works.

Thus, how do the outcomes of the analytical model compare to proponents' claims that torture is used only against knowledgeable detainees who refuse to provide information and that once it is used, (the threat of) torture generates all, or nearly all, the valuable information possessed by those knowledgeable detainees? How do the outcomes compare to their claims that torture can be limited and regulated so that it will not be used against cooperating detainees who have provided all their information, nor against innocent detainees, nor exceed the minimum frequency and severity 'necessary' to elicit information from knowledgeable detainees?

The torture of innocents is inevitable
Our first four outcomes above (*Complete valuable information, surprise torture, Incomplete valuable information, surprise torture, Complete valuable information, selective torture, and Incomplete valuable information, selective torture*) might seem to provide support for proponents, since at least some valuable information is provided in each outcome. Before considering information, notice first, however, that that both outcomes require the interrogator to be willing to torture an innocent detainee. The interrogator must be willing to torture any detainee who fails to provide information—including an innocent detainee—in order to credibly threaten a cooperative detainee into

⁷ In the case of the innocent detainee, of course, he cannot cooperate if the questions are objective and "cooperating" under leading questioning means falsely confirming what the interrogator wants to hear.

revealing information. Unless an interrogator actually tortures after failing to receive valuable information, according to the proponents' own rationale, no cooperative detainee will reveal that information. In other words, the only way to get any valuable information (just how much we examine shortly) is to violate the necessary condition prohibiting the torture of innocents.

This also, however, rules out the *Torture Justification Outcome*, which claims it is possible to get valuable information without torturing innocents. It is not. The only way to even make it possible to get any valuable information is to torture innocents; refusing to torture innocents means no valuable information. This necessary and inescapable truth explains why the history of torture is littered with the blood and pain of innocents.

Information unreliability

Although we have already found one normative condition and the torture justification outcome to have been violated, return to the four outcomes in which valuable information is provided by a cooperative detainee: *Complete valuable information, surprise torture, Incomplete valuable information, surprise torture, Complete valuable information, selective torture, and Incomplete valuable information, selective torture*. There are two problems for information reliability here. First, even if we count all four outcomes, then a maximum of exactly one-half of all possible outcomes result in valuable information. It is highly unlikely you would call a surgical procedure, an airplane, or your car reliable if it worked as often as a coin flip came up heads. Second, the problem is actually worse because the incomplete disclosure versions of the two outcomes mean a detainee got away with precisely that which torture was supposed to prevent: information hiding.

Thus, we are left with two of the eight outcomes, *complete valuable information, surprise torture* and *complete valuable information, selective torture*, and so the ratio drops to one in four - unreliable indeed.

It is possible to assess the reliability of information and later the extent of torture more visually. As we saw above, the outcomes depend on three crucial beliefs: the detainee's information revelation threshold (the point at which the cooperative detainee is willing to divulge information because he believes he won't be tortured [more] if he does so), the interrogator's belief about whether the detainee is hiding more information or not (information hiding threshold), and the interrogator's belief about whether a detainee who provides no information is innocent or not (innocent detainee recognition threshold). Since these are all beliefs, or probabilities, each can be arrayed on an axis from zero to one. Combining them creates a three dimensional cube, as in Figure 1.

Since the cube captures the full range of all three beliefs, all of the eight outcomes inhabit this space. The entire cube represents the universe of what can happen in the model. Any particular subset of that space defined by different combinations of being above or below the thresholds marked with dotted lines is an outcome or set of outcomes. The greater the volume taken up by an outcome, the more likely it is because it is supported by a greater range of beliefs. This allows us to compare the relative space taken up by different outcomes to the total possible space and so an outcome's relative likelihood. In particular, we will examine how much of this space is taken up by the two outcomes with complete valuable information to see how likely they are. Later we will examine how much of this space is taken up by torture to see how likely it is.

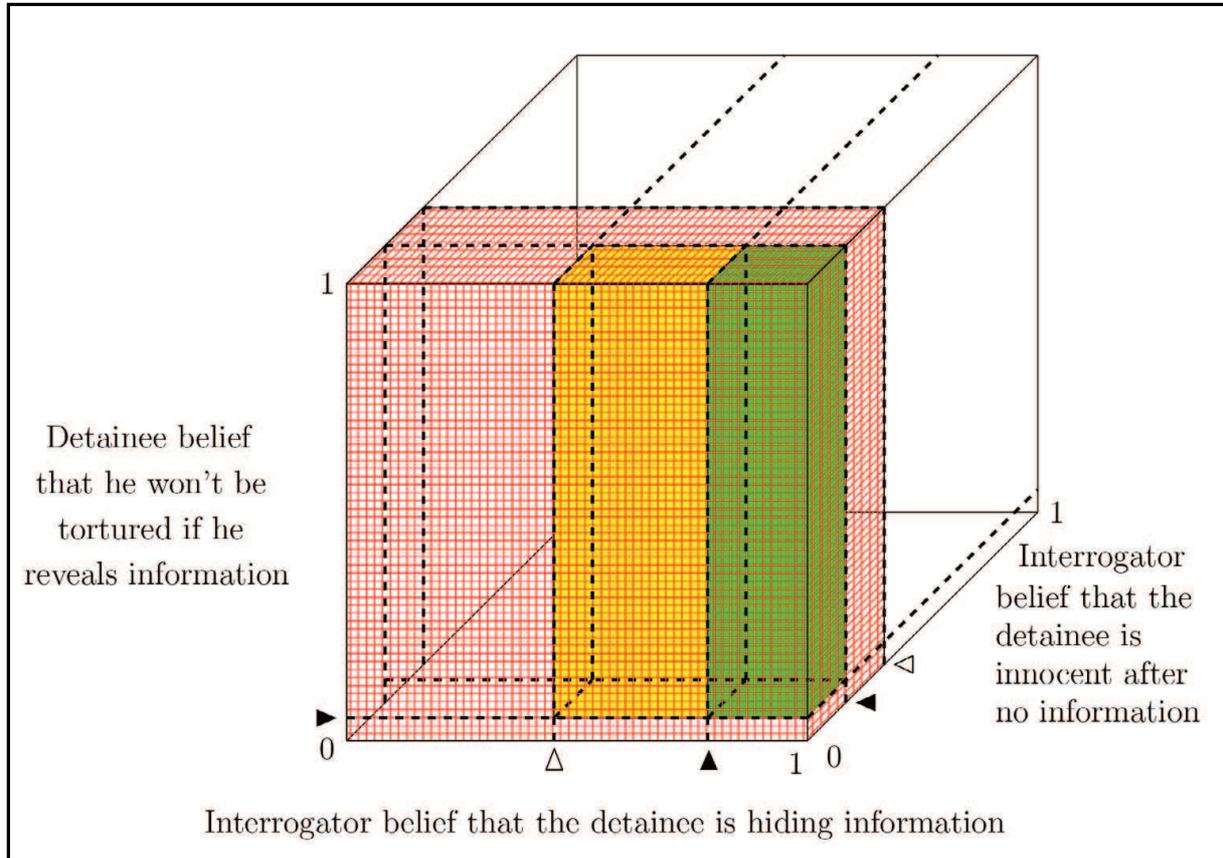


Figure 1: *Maximum extent of valuable information and torture*

The region highlighted in yellow represents the *Complete valuable information, surprise torture outcome* and the region in green encloses the *Complete valuable information, selective torture outcome*, both under objective questioning. Both lie above the detainee's information revelation threshold on the left vertical axis (which is why he provides information) and below the interrogator's innocent detainee recognition threshold marked by the solid triangle on the right diagonal axis (the interrogator believes that if she fails to get information it's because the detainee is lying, not because he's innocent). The line separating the two outcomes marked by the solid triangle on the horizontal axis is the interrogator's information hiding threshold. The *Complete valuable information, selective*

torture outcome in green lies to the right of this plane because the interrogator believes the detainee has given up all his information and so does not torture him. Note, however, that in this outcome the interrogator would torture an innocent detainee for not providing information because, although the detainee is truly innocent, the interrogator falsely believes he is not and she is committed to torturing a detainee who fails to reveal information. This is the reason for the "selective" in the name of the outcome. If the interrogator does not believe the detainee has revealed everything (i.e. for values on the horizontal axis to the left of the solid triangle) she tortures. The detainee, however, thinks that the interrogator believes he has divulged everything and so she will not torture him. This difference in

beliefs is represented as the region along the horizontal axis between the empty and solid triangles. If so, then he will provide information and then be surprised when she tortures him afterward, generating the *Complete valuable information, surprise torture* outcome captured by the region in yellow. Including the surprise torture outcome is generous to proponents since torturing a detainee after providing full information violates the normative model's prohibition on torturing detainees 'unnecessarily' and so counts against the proponent model. Thus an interpretation more consistent with proponents' own promises to limit and control torture would slice off the yellow region represented by the surprise torture outcome, leaving only the green region, which itself violates the prohibition on torturing innocents.

Even though both outcomes violate the proponent model's restrictions on unnecessary torture, what can we say about the relative size of the regions with valuable information in yellow and green? If they took up the entire cube that would mean torture was very reliable because you would get valuable information for a wide range of beliefs on all three axes; high, low, and everything in between. Conversely consider the other (absurd) extreme and imagine the outcomes collapsed to a single point or dot in the cube because valuable information only occurred for three specific values on each of the three axes, say .34, .67 and .41. This would obviously reduce the reliability of torture quite a lot. It turns out that the proponent's own logic provides some constraints on how large this area can be.

First, consider the interrogator's information hiding threshold marked by the solid triangle along the horizontal axis. In assuming that interrogators prefer not to torture if they can get the information they

want, we build in costs to them of using torture. This is why it is possible that they would choose not to torture if they thought a detainee was innocent or they thought they had received all of a cooperative detainee's information: they would bear the cost of torture with no compensating information benefit. If, as is the case in real life, these costs are very low, however, this pushes that threshold to the right so that it is past one-half. After all, an interrogator won't worry too much about continuing to torture a cooperative detainee past the point he has provided everything he knows if her costs of using torture are low. This means, though, that the green region shrinks to the right, making it smaller and so the outcome less likely.

Now consider the interrogator's innocent detainee recognition threshold marked by the solid triangle along the diagonal axis on the lower right. Imagine for a moment you are the interrogator. Which would you think was the most likely of the three theoretically possible types of detainees—knowledgeable and cooperative, knowledgeable and resistant, or innocent? How probable is it for you that the naked, hooded, shivering, semi-starved man shackled to the ceiling in front of you is completely innocent? You will not think this is very likely. You might consider it somewhat more probable that a detainee is knowledgeable but resistant—at least so far. You are not getting anything out of him but you have confidence that your techniques will work eventually. The upshot here is that you will think the most likely type of detainee shackled in front of you will break under your torture (i.e. is cooperative) and it is far less likely that he has information but can never be broken (resistant), and finally even less likely that he is completely innocent. That, however, pushes the threshold under one-half and it

gets closer to zero (to the front of the cube) as the probability of a resistant detainee gets closer to zero. Once again this has the effect of shrinking the green and yellow regions with valuable information by squeezing them toward the front of the cube from the back.

Putting these two thresholds together tells us something about information reliability before even considering the last threshold on the vertical axis. Squeezed to the front, right corner, the green *Complete valuable information, selective torture* outcome cannot occupy even a quarter of the total volume of the cube. In other words, you can expect to get valuable information without ‘unnecessarily’ torturing a cooperative detainee fewer than one out of four times you use torture (and only if you are willing to torture an innocent detainee). An additional willingness to violate the prohibition against unnecessary torture and torture detainees beyond the point they have any more information (i.e. including the *Complete valuable information, surprise torture* outcome in the yellow region) pushes the total volume with valuable information to the left, but only a little, just under one-half on the horizontal axis, a marginal increase in reliability.

Finally, take a look at the last threshold, the cooperative detainee’s information revelation threshold on the left vertical axis. For values above this threshold, the detainee reveals information (that is why the yellow and green regions are bounded on the bottom by this threshold). For values below it, the cooperative detainee refuses to reveal information because he thinks he will be tortured anyway. Consistent with the logic of torture, the detainee’s threshold for revealing information is the ratio of the value of his information to the pain of the torture. Once pain becomes the method of interrogation and all else being equal, a detainee is less

likely to give up more valuable information than less valuable information and, consistent with the (empirically flawed) assumptions behind using torture in the first place, more pain is expected to elicit more information than less pain.⁸

Unlike the other two thresholds, since this one is determined by the simple ratio of information value to pain, there is no general constraint on its location. To be generous to proponents, Figure 1 pushes the threshold all the way down near zero in order to maximize the size of the yellow and green valuable information regions. Even so, however, it is clear that valuable information takes up a relatively small proportion of the total space, making it an unlikely and so unreliable outcome from torture. Moreover, the simple ratio of information value to torture means that in order for the information revelation threshold to be so far down (making the yellow and green regions larger and so those outcomes more likely), either the value of the information must be very low or, if the the information value is high, then the torture must be exceedingly brutal to compensate and still make the detainee reveal it.

Finally, it is important to consider what the much larger, non-shaded area to the rear of the cube signifies: (1) truthful, accurate, but non-valuable information, (2) false and misleading information, and (3) no information whatsoever, all from a cooperative detainee who has valuable information. The latter two are particularly problematic for the pragmatic argument for

⁸ Flawed because, in fact, more pain may very well *reduce* the value of any information provided (O’Mara, 2015, pp 48, 106-107, 115-132; Pérez-Sales, 2016, pp 187-188, 222, 267-271). I thank Pau Pérez-Sales for enjoining me to clarify this point.

torture. Proponents rarely factor in the costs of misleading information, of time, money, and resources diverted from actual to non-existent threats or sources of information. Nor is the absence of information to be dismissed as merely neutral, as neither benefit nor cost. Torturing someone who provides no information also costs time and resources in addition to being ultimately ‘unnecessary’ (since it resulted in no information). Moreover, this is precisely what is not supposed to happen once torture is introduced into interrogations; torture is the method that is supposed to ‘break’ resistant detainees, to compel information they would otherwise not release.

Slippery slope 1: Torture will be frequent

Proponents of interrogational torture claim that the information will be valuable while ‘necessary’ torture is kept to a minimum. The total frequency of torture will be low, cooperative detainees will not be tortured long after they have provided all their information, nor innocent detainees tortured for telling the truth. We have already seen that the first part about information is false, as is the claim about innocents and cooperative detainees. What about the total frequency of torture?

Once again, we can assess this claim in two ways, by looking at the proportion of outcomes in which there is ‘unnecessary’ torture as well as the likelihood of torture represented by the cube in Figure 1. Of the eight possible outcomes, there is torture of one or more types of detainees in seven of them. We have already seen that even when valuable information is provided by cooperative detainees, innocent detainees are tortured. In the surprise torture outcome a cooperative detainee is tortured after providing information, as is an innocent for telling the truth. In the *False confirmation*,

selective torture outcome, the cooperative detainee is tortured for refusing to confirm a leading question. And resistant detainees are tortured in all seven for refusing to provide valuable information. From this perspective, then, torture will be frequent, not infrequent.

Turning to the second, visual, perspective, notice the red-hatched region in Figure 1. This region, in front of the interrogator’s innocent detainee recognition threshold marked by the empty triangle on the lower right diagonal axis, represents ‘unjustified torture’ in one form or another: the torture of a cooperative detainee after providing information, the torture of an innocent detainee for telling the truth, or both. The complementary clear volume to the rear of the open arrow on the lower right represents the *No information, no torture* outcome in which no detainee is tortured. Assuming that the costs to the interrogator and the state of using torture are low, as seems reasonable, this hatched volume is just under half of the entire cube, meaning that even when torture is introduced as a supposed last resort, it will be used about half of the time under the assumption most generous to proponents—hardly limited torture. Moreover, this depends on the generous assumption that interrogators are willing to forgo torture if they think a detainee is innocent when they do not get valuable information from them (to the rear of the empty arrow on the lower right axis). Less generously but more realistically, if we think that in most cases an interrogator is likely to interpret a failure to elicit valuable information as a sign of resistance, rather than innocence, then the entire space in front of the empty arrow, where all the remaining seven outcomes are located, is taken up by torture. Either way, it seems clear, torture will be far more frequent than its proponents imagine and claim in their justification of the practice.

Slippery slope 2: Torture will be brutal

In addition to claiming that the frequency of torture across detainees can be minimized and controlled, proponents also argue that the severity or brutality or intensity of torture visited upon any one detainee can likewise be minimized and controlled to roughly that ‘necessary’ to elicit information. This, however, conflicts with the proponents’ own basic pain–information logic. Recalling that the cooperative detainee’s decision to reveal information is based on the ratio of the value of the information to the costs of the torture, consider the incentives for the interrogator. She cannot control the value of the information possessed by the detainee and she cannot know the effects of any particular torture (some detainees can withstand cramped confinement longer than others). All she can control is the severity or degree of torture she inflicts on a particular detainee. Once torture becomes the method of interrogation, her only hope of increasing the likelihood of getting information is to increase the degree of pain, to torture more. Once again it may be easier to see this visually by returning to Figure 1. Grabbing the solid arrow on the vertical left axis and driving it down is the graphical equivalent of what was just said about the incentive facing the interrogators. Driving down the threshold means increasing the ratio of pain to information in an effort to push him over his threshold; that is the only tool she has left once she starts torturing. The point is that, even discounting the very real and significant effects (1) of a sincere motivation to ‘get the bad guys,’ (2) of an organizational culturing pushing interrogators to ‘do what it takes,’ (3) of pressure from higher-ups, and (4) of psychological biases, the incentives *alone*

from the logic of interrogational torture commit an interrogator to increasing pain, to driving down that threshold in the hope of making it more likely a detainee will talk. In other words, here too there is a slippery slope, violating proponent limits on torture’s brutality.

Conclusion

There is a basic contradiction between the unavoidable premise behind the age-old logic of torture—more torture means more information—on the one hand, and the claims of those who defend a program rooted in this very same logic—torture can be minimized yet information maximized on the other hand. We have seen that neither is true. The only thing reliably effective about interrogational torture is its ability to generate slippery slopes of frequency and brutality, to escape the limits and controls imposed on it. Torture is maximized yet information is minimal. These results emerge from the proponents’ own model for how torture is supposed to work, including limits on the torture ‘necessary’ to elicit information.

By tracing out the different possible combinations of detainees and uncertainties associated with interrogational torture, it is possible to generate the wide range of outcomes that we observe in the real world. This systematic approach also helps demonstrate why the tendency to justify interrogational torture as a one-off event with reference to the ticking bomb scenario, whether by utilitarian philosophers or politicians, is unhelpful. When states sanction torture, it becomes institutionalized, it becomes a system and bureaucracy of torture. A rigorous examination of interrogational torture in even a simple model complements other approaches by demonstrating how interrogational torture twists logic as well as bodies.

Acknowledgments

I am grateful to Carol Rounds, Robert Houle, Pau Pérez-Sales, and two anonymous reviewers for their comments on earlier versions of this paper. Remaining errors are, of course, my own.

References

- Améry, J. (2009). *At the Mind's Limits: Contemplations by a Survivor on Auschwitz and its Realities*. (T. Rosenfeld, S. & Rosenfeld, S.P., trans.). Bloomington, IN: Indiana University Press.
- Baliga, S., & Ely, J. C. (2016). Torture and the commitment problem. *Review of Economic Studies*, 83(4), 1406–1439. <https://doi.org/10.1093/restud/rdv057>
- Denbeaux, M., Denbeaux, J., Gratz, D., Gregorek, J., Darby, M., Edwards, S., Hartman, S., Mann, D., & Skinner, H. (2006). "Report On Guantanamo Detainees: A Profile of 517 Detainees through Analysis of Department of Defense Data." Seton Hall University School of Law.
- Elbert, T., Rockstroh, B., Kolassa, IT., Schauer, M., & Neuner, F. (2006). "The influence of organized violence and terror on brain and mind: A co-constructive perspective". In *Lifespan development and the brain: The perspective of bio-cultural co-constructivism* (Baltes, P. B., Reuter-Lorenz, P. A., & Rösler, F., eds). Cambridge: Cambridge University Press pp. 326–375. <https://doi.org/10.1017/CBO9780511499722>
- Elbert, T., Schauer, M., Ruf, M., Weierstall, R., Neuner, F., Rockstroh, B., Junghofer, M. (2011). *Zeitschrift Für Psychologie*, 219(3), 167–174.
- Gibbons, R. (1992). *Game Theory for Applied Economists. Comparative and General Pharmacology* (Vol. 34). <http://doi.org/10.1017/CBO9780511791307.017>
- ICRC. (2016). *People on War: Perspectives from 16 Countries*. International Committee of the Red Cross. Geneva, Switzerland.
- Johnson, D. A., Mora, A., & Schmidt, A. (2016). The strategic costs of torture: How "enhanced interrogation" Hurt America. *Foreign Affairs*. 95(5):121-132
- Johnson, J. (2016). Trump says torture works, backs waterboarding and much worse. *Washington Post*. February 17
- Karpinski, J. (2005). Interview, Janis Karpinski, "Frontline: The Torture Question". *pbs.org*. October 18, 2005. Retrieved from: <http://www.pbs.org/wgbh/pages/frontline/torture/interviews/karpinski.html>.
- Maercker, A., & Forstmeier, S. (2011). Torture and atrocity sequelae in methodologically sophisticated studies. *Zeitschrift Für Psychologie / Journal of Psychology*, 219(3), 182–183. <http://doi.org/10.1027/2151-2604/a000066>
- New York Times, The*. (2017). "The Guantánamo Docket." Retrieved from: <https://www.nytimes.com/interactive/projects/guantanamo/detainees?mcubz=1>.
- O'Mara, S. (2009). Torturing the brain. On the folk psychology and folk neurobiology motivating "enhanced and coercive interrogation techniques." *Trends in Cognitive Sciences*, 13(12), 497–500. <http://doi.org/10.1016/j.tics.2009.09.001>
- O'Mara, S. (2011). On the imposition of torture, an extreme stressor state, to extract information from memory: A baleful consequence of folk cognitive neurobiology. *Zeitschrift Für Psychologie / Journal of Psychology*, 219(3), 159–166. <http://doi.org/10.1027/2151-2604/a000063>
- O'Mara, S. (2015). *Why Torture Doesn't Work: The Neuroscience of Interrogation*. Cambridge, MA: Harvard University Press.
- Pérez-Sales, P. (2016). *Psychological Torture: Definition, Evaluation, and Measurement*. Philadelphia: Taylor & Francis Group. doi. 10.1097/WTF.0000000000000152
- Rejali, D.M. (2007). *Torture and Democracy*. Princeton, NJ: Princeton University Press.
- Rumney, P.N.S. (2014). *Torturing terrorists: Exploring the limits of law, human rights and academic freedom*. London: Routledge.
- Schiemann, J.W. (2016). *Does Torture Work?* New York: Oxford University Press.
- Szymborska, W. (1998). *Poems new and collected, 1957–1997*. (Stanislaw Baranczak and Clare Cavanagh, trans.) New York: Harcourt Brace & Co.
- United States Senate. (2014). *Select Committee on Intelligence Study of the Central Intelligence Agency's Detention and Interrogation Program, Findings and Conclusions, Executive Summary*. Washington D.C.: United States Senate. Approved December 13, 2012, Updated for Release April 3, 2014, Declassification Revisions December 3, 2014. Retrieved from: <http://www.intelligence.senate.gov/study2014/sscistudy1.pdf>.

Related to 'Tortured logic: Information and brutality in interrogation' by John W. Schiemann. (pages 64 to 78)

Comment I

Old sins cast long shadows: Further reflections on the power of rapport over retributive Interrogation.

Laurence Alison*, Emily Alison*

Schiemann's article is a refreshing and welcome take on the arguments for and against torture. The notion of using a dendritic set of outcome options based on inputs that relate to reliability, sadism, limits and control is a useful way to approach the problem. His argument that, taking all these features into account results in the logical outcome that it should never be used is a compelling one. However, we do think, nuanced though that argument is, other things need to be taken into account. For too long the scientific debate has been polarised into one of moral outrage at even contemplating 'harsh' methods without providing an alternative solution. More recently, evidence has been found for 'rapport based' methods that supply a viable, ethical strategy (Alison and Alison, 2017). To be clear, our view is that, on moral grounds, it is repugnant and should never

be contemplated. However, as scientists, we also need to consider the scientific merits of any proposed 'method' for extracting information. There is a moral argument and then there is a scientific argument and too often the two have been conflated. So, Schiemann's logical, decision tree analysis is very welcome.

However, there are some additional points we would wish to make. The Mitchell and Jessen post 9/11 enhanced interrogation techniques are an important one to consider (Mitchell and Harlow, 2016). For example, with regards to Mitchell and Jessen's 'methods' they claim that they did not extract information from Khalid Sheik Mohammed during the waterboarding sessions but rather after them, when he was asked to consider giving information to avoid such an experience happening again. Indeed, Mitchell claimed to be a strong proponent of rapport based methods, though he regularly uses dismissive terms such as 'tea and sympathy' to describe them. His argument was that inducing learned helplessness could encourage a detainee to cooperate and it was after inducing this state, through the enhanced interrogation techniques that the detainee would then be more amenable to the 'rapport' based methods. Thus, the argument was not that torture worked 'in the moment that it was being used' but that, effectively, it 'softened' the detainee up in order to presumably become more 'susceptible' to the rapport based methods.

In order to really understand the efficacy of torture techniques, we do need to engage with and thoroughly consider the arguments put forward by torture proponents. In considering whether there are scientific arguments that support torture, it does not mean we lose the ability to see the immoral reasons why torture should not be used. Instead, we need to fully consider

*) Critical and Major Incident Psychology, Psychological Sciences, University of Liverpool.
Correspondence to: L.J.Alison@liverpool.ac.uk

the counter arguments that might be put forward by torture proponents. These include the following (not exhaustive) list:

- If it doesn't work, why do we find ourselves repeatedly using it?
- Torture is not expected to work at the time of doing it, but rather to soften the detainee up to make him/ her more amenable to rapport based methods
- As a general rule, we should not torture but there may be some occasions (e.g. ticking time-bomb argument), in some desperate situations and with certain types of detainee where it might be necessary.

The first question is an interesting one. Alison and Alison (2017) and Carlsmith and Sood (2009) have argued that one motivation is surely retribution with no real ambition to seek information at all. This notion emerged as long ago in the 16th century to justify the trial and execution of men, women and even children for witchcraft. In many cases, the principle purpose was to rid communities of problematic individuals (Gaskill, 2005). In essence then the end goal is not information but publically condoned retribution. One cannot help think that during times of national crisis a key motivating force is to exact revenge. Sternberg (2003) has carefully considered the role of hate as a specifically directed emotion that begins with (i) generation of hatred and disgust at the target group, (ii) anger and a sense of 'them and us' (iii) contempt and, finally (iv) punishment for those that do not support the in group. In such cases torture does 'work' to extract reliable intelligence or information—it works to enact retribution on those who have wronged. The interrogators become the vehicle through which we make that retributive act.

The second statement argues that the torture itself didn't work other than to show that these phases can be stopped and to then

offer an 'out' to a 'softened' up detainee. We think this argument, although marginally more nuanced than the over simplified model that some of the anti-torture camp have caricatured it as, still fails to hold any logic. The argument seems to be that what an interrogator offers the detainee is the following, 'this doesn't have to happen again—have a think about whether you want to tell us about X and if you do, you then won't have to be subjected to this again'. Thus, the notion is of torture + threat + contemplation period = consideration that it may be best to talk. This avoidance of aversive stimuli in exchange for information is, in our view also highly problematic. First, and as per Schiemann's very sensible train of thought, what guarantees does the detainee have that it will indeed stop if information is forthcoming? Second, why not lie about the information in order to avoid giving up key information? Third, even if the detainee tells the truth will s/he be believed and if not will the torture resume? Fourth, even if the detainee tells the truth, what happens if more information or more details are required? In such an instance, do we go back to the pattern of torture + threat + contemplation? These detailed other decision paths need to be considered and, indeed, we are in the process of doing so (Surmon Bohr, Alison and Alison, *In Preparation*, 2018). However, our analysis always results in the same outcome—namely a zero-sum game in which no one wins, since neither the detainee, nor the torturer can know in advance and for certain what each other's' next move will be.

The final statement suggests a greater degree of reserve in the use of torture. Namely, it intimates that it should only ever be used as a final resort when the stakes are high and all else has failed. Tempting though that may be (especially in desperate times), it offers only a short term solution (at best)

and even then, one where the reliability of each piece of ‘life saving intelligence’ is unknown. In Cobain’s (2013) very eloquent book ‘Cruel Britannia’—an extensive coverage of the history of the use of torture by the British, he does allude to the notion that some (and note the word ‘some’) information may have been extracted through torture (though the book is strongly anti torture in its general tenor). For example, he notes some information did come from detainees in the second world war London Cage, the five techniques used in the Mau Mau uprising as well as in the Troubles in Northern Ireland in the 1970s. However, there are two separate, but critically important counter arguments that still question the argument of ‘sometimes and only in extreme circumstances’.

The first is a very simple one: where is the evidence? Although several books, authors and interrogators indicate their torture based methods worked, our request remains, ‘show me’. We have been fortunate enough to observe countless hours of field interrogations and all we have ever seen is rapport working. Even mild sarcasm has a deleterious effect (Alison, Alison, Noone, Elntib and Christiansen (2013)). We are quite prepared to adjust our rapport based model if we can be furnished with a sample of audio or video clips showing enhanced interrogations working but the data from those that espouse and defend those methods has never been forthcoming. It may be hard for us to submit to evidence to the effect of torture on moral grounds but if we see it working we will say it works. No one has been in a position to say this so on those basic most fundamental principles of ‘convince me by showing me’ - there is nothing. The second point relates to what we call the ‘long interrogation game’. There are countless examples of the consequences of torture as a recruiting tool for extremists, a massive blow

to an organisation and country’s reputation, a transgressive act from which the interrogator may never recover and terrible psychological and physical consequences for the detainee. The adage of ‘old sins cast long shadows’ is especially relevant to the history of torture and so, whilst we accede to the notion that through rapport you cannot always make things better, with torture, you most certainly can make things worse.

References

- Alison, L. & Alison, E. (2017). Revenge Versus Rapport: Interrogation, Terrorism, and Torture. *American Psychologist*, 72(3) 266-277. DOI 10.1037/amp0000064.
- Alison, L., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why Tough Tactics Fail and Rapport Gets Results: Observing Rapport-Based Interpersonal Techniques (ORBIT) TO Generate Useful Information from Terrorists. *Psychology, Public Policy and Law*, 19, 411-431. DOI 10.1177/0093854815604179
- Cobain, I. (2013). *Cruel Britannia: A Secret History of Torture*. Portobello Books.
- Carlsmith, K. & Sood, A. (2009). The Fine Line Between Interrogation and Retribution. *Journal of Experimental Social Psychology*, 45, 191-196. DOI 10.1016/j.jesp.2008.08.025
- Gaskill, M. (2005). *Witchfinder: A Seventeenth-Century English Tragedy*. London, United Kingdom: John Murray.
- Mitchell, J. E. & Harlow, B. (2016). *Enhanced Interrogation: Inside the Minds and Motives of the Islamic Terrorists Trying to Destroy America*. Penguin Random House.

Comment II

Glenn L. Carle*

For nearly twenty years now, I have spoken of my involvement in the “Enhanced Interrogation Techniques” (EIT) Program,

*) former National Intelligence Officer for Transnational Threats, career CIA officer and interrogator of a High Value Detainee

trying to show the American public how the United States unnecessarily betrayed its values, deluded and lied to itself about what it did, and changed its culture and institutions for the worse. Atavistic demons of torture possessed us and made us a cruder society, and a spurious ticking time bomb scenario seduced us to betray our values.

John Schiemann's *Tortured Logic: Information and brutality in interrogations* coolly and rationally demonstrates to us how the EIT program is founded on faulty assumptions, and how the arguments of its proponents themselves demonstrate that the program's premises guarantee that torture become frequent and progressively brutal, once accepted as a "useful" tool of interrogation. Schiemann shows this through a game theory analysis of possible behaviors by detainees and interrogators. Game theory was far from our minds when I was interrogating my detainee, and yet Schiemann's analysis of the motives and calculated actions of detainee and interrogator, acting within a framework calling for torture in various circumstances of cooperation or resistance by a detainee, brought the tragic pressures my colleagues and I dealt with painfully back to life. Schiemann accurately captures the nearly irresistible dynamic that leads to torture once "EIT"s have been accepted as a legitimate technique, and he dispatches the falsehoods used to assert that torture can extract useful information and yet be limited in extent once begun. It does not, and it cannot.

I note with tragic irony that, literally, Schiemann engages in cooler, more, and deeper analysis of the rationales, costs, and benefits (sic) of "enhanced interrogation"—torture—than occurred in the entire CIA throughout the Enhanced Interrogation Program. This devastating point also emerges in the historically important,

and utterly accurate, Senate Intelligence Committee Report on Torture. My one surprise in reading the Senate Report was to discover how widely-shared was opposition to the premises and practices of the EIT program among those of us involved in it, as was the conclusion that EITs were ineffective, illegal, counter-productive, and mindless. Schiemann's conclusions do not surprise me, as they mirror those I reached as I struggled with how to conduct an honorable, legal, effective interrogation. His analysis, however, surprises by so easily tearing down the Potemkin Village rationale of those defending the need for and utility of torture. But Schiemann brings reason to an argument of faith; the cerebral cortex and reason are so often powerless confronted with the amygdala and atavism. His logical proofs will not change the minds of torture's defenders. It should, however, shift the trajectory of the debate going forward for those who will respond to tomorrow's existential moments.

Schiemann examines eight possible outcomes in the dynamic between interrogator and detainee in an "EIT" program. The outcomes range from the detainee providing completely valuable information without torture, or from it; to providing incomplete but valuable information without torture, or from it; to ambiguous information, falsely confirmed information, no information... Each of the eight outcomes is a function of whether torture was used extensively, selectively, or not at all. The conclusions of the analysis are simple and, in my experience, spot on. When torture is part of an interrogation doctrine:

- 1) The torture of innocents is inevitable
- 2) The information provided is unreliable
- 3) Torture will be frequent and increase to universal use, despite the claims of the

EIT program's proponents that it would only be used when necessary

4) Torture will be progressively brutal

The game theory analysis of whether to torture, what to do in the event of one response by the detainee or another, however, is more thoughtful than was the practice in fact. Detainees were assumed to be "guilty"—to have critical information—because they were detained. Therefore, EITs would be used as a matter of course. Otherwise, the detainee would not have been brought into the detainee, EIT program. Guilt and innocence were not part of any interrogation equation. Further, I often struggled against the explicit "guidance" that the lack of an answer by the detainee on a given question "proved" that he was withholding information, and that, therefore, I had to increase "pressure" on the detainee so that he would provide the information he was withholding. My retort that the detainee's lack of answers to specific questions, in my assessment, often (not always) indicated that the detainee did not know the answers, and was responding truthfully, was taken as proof that I had "fallen in love" with my detainee; I should focus, instead, on forcing the detainee to provide the information his lack of answer proved he was withholding. My further protestations that this response was stupid and irrational (something like "asserting that his professed lack of knowledge proves he is withholding information does not, ipso facto, prove that the detainee knows the information and is withholding it. I assess that he answered truthfully...") simply tarred me as an apostate. I have pointed out elsewhere that the typical response of clergy to such jarring doctrinal challenge is to burn the apostate, not re-examine the doctrine.

This recurring argument with the EIT program's masters raises a key dynamic that Schiemann's game theory analysis, and his assessments, underemphasize. It can only be so, since Schiemann is rational and analytical (and his assessments accurate.) Kafka was far more important to the genesis and running of the EIT program than was any analysis of whether EITs—torture—made any sense. In *The Trial* Kafka captures the dynamic that obtained in the EIT program: "No," said the priest. "It is not necessary to accept everything as true, one must only accept it as necessary." CIA's—the US Government's—EIT program and, indeed, much of the US Government's "War on Terror," was atavistic and visceral, not a rational response to a terrorist threat, or to how to conduct a successful interrogation.

The reasoning among many was that the nation was under threat, the institution had decided to engage in EITs, the government stated that EITs were not torture, so therefore they were not torture. Assessing varied outcomes of interrogation, under various assumptions by detainee and interrogator—this was far beyond any prepared approach, or subsequent analysis, of whether to engage in EITs or whether they worked. The program was the program, detainees had information because they were detainees, and therefore one would use EITs on them. As I was literally told once when I was trying to place a specific terrorist group's actions in a regional context: "Sociology is great, Glenn. But this is not a graduate school seminar. Just find me the f---ing terrorists."

There was another defining component to the EIT program: The rightly infamous "one percent doctrine." The one percent doctrine, evinced by Vice President Cheney, in the context of counter-terrorism operations and interrogations, held that

the US must act as though the individual were guilty, even if there were only a “one percent” chance that an individual was what we apprehended might be the case. I saw, I heard, this doctrine invoked as our operative guidance over and over, sometimes sincerely, sometimes as what I took to be a thin rationalization to exculpate one party or another. Given this, it was certain that any detainee considered an HVT, or a threat (which, by definition, a detainee was) would be subjected to EITs. Careful assessments of game theory possibilities were beyond the reality we all lived. A detainee was part of the one percent, surely, and surely would be subject to EITs.

Schiemann rightly debunks the assertions of the EIT program’s proponents: That torture gives reliable information and that torture can be controlled. He does not note, however, two important and telling points: First, the only “proponents” of the EIT program are the Bush administration and CIA officials involved in the program’s conception and management and, second, all supposed defenses of the program’s efficacy and success have been made *ex post facto*. These views were not strongly advocated at the time, but have been made since to justify a program that clearly violates numerous American and international laws—whatever the ratiocinations of the political hacks who drafted the infamous “torture memo” which provided legal cover for the program (“We’re okay. We’re covered” I was told at the time, when I asked, among other things, “But what about the Geneva Conventions?”...)

Proponents repeatedly cite the mantra of the “twenty” terrorist operations that information obtained from EITs enabled the CIA and FBI to stop. But, without going into classified information, it has been shown that these twenty terrorist operations (sic) either were stopped without EIT-

provided information; or more frequently were vastly overblown in how they were presented to policymakers (it should be noted that both senior CIA officers and senior Bush administration officials were frequently sold, frankly, a false, although often sincerely held, narrative by the counter-terrorism bureaucracy of the CIA. But that is a separate, albeit critical, story.) To my knowledge, none of the twenty terrorist operations stopped by the CIA and FBI had anything to do with information provided by EITs, and I was in as good a position as most anyone to know what we knew, and how we knew it. As I repeated and repeated during my years working on the “War on Terror”: “The closer one looks, the less one sees.” And, the closer one looks at EIT-produced information, the more one sees that Schiemann’s assessment gets it right: torture provides bad information, to which I add: torture also perverts the US as a nation, replacing law and reasoned action in the national defense, with an atavistic series of assumptions and actions, which harm us more than protect us.

Game theory can seem bloodless to those studying international relations. Perhaps. But, in this instance, at least, it has exposed the fraud of the defenses made of EITs—of torture. If we take Schiemann’s insights to heart, we can undo some of the harm the EIT program—we—have done to ourselves.

Comment III

Hans Draminsky Petersen, MD.*

I should start with the admission that I have little or no knowledge of game theory and

*) ex-member and vice-chair of the UN Subcommittee on Prevention of Torture (SPT)

this may well be the reason why I find some elements of John W. Schiemann's article *Tortured logic: Information and brutality in interrogation* illogical. However, since it can be assumed that most readers of the *Torture Journal* lack knowledge about game theory, my observations and response to the article may be relevant.

The model described in the paper only deals with torture scenarios where the objective is to obtain key intelligence information and not with the vast majority of torture cases where obtaining confessions from criminal suspects is the objective - although that problematic is mentioned briefly and put into the context of the model.

The model works with the three dimensions of a cube (Figure 1). Each of the three axes have thresholds: the detainee's threshold for revealing information; the interrogator's threshold for believing that the detainee is hiding information; and the interrogator's threshold for believing that the detainee is innocent. Together these three thresholds demarcate sections of the cube, which are described as outcomes of interrogations; obtaining from the detainee all or part of the desired information as replies to leading or open-ended questions in conjunction with torture, leading to more torture or to the end of torture. The author explains that by moving the thresholds and thereby changing the sizes of the segments, the likelihood of obtaining reliable information without having to use torture unnecessarily can be estimated.

Some reservations as to the logic of the model: One of the dimensions in the figure deals with the interrogator's belief that the detainee is hiding information. A threshold for the interrogator is indicated; to the left the detainee is believed to retain information, which consequently calls for torture. I infer that to the right of the

threshold the detainee is believed to not retain information and consequently torture should not be used. This, however, is not the case since the green segment to the right of the threshold (Figure 1) includes innocent detainees who cannot give information and therefore are tortured. I find this in conflict with the concept of a threshold.

In the text it is explained that the rear segment in the figure includes cooperative detainees who give no information whatsoever. I cannot understand how such a detainee can be classified as cooperative. These issues therefore put me in doubt of whether I have understood the model correctly.

Conceptual reservations: It is assumed that the victim is in a constant dialogue with him or herself assessing all the time how far s/he will continue without giving all the information s/he has, while realising that the price for not giving the required information will lead to continued torture and while assuming that giving information - the whole or parts of it - will mean an end to torture. (S/he may well be disappointed to see that torture continues after having given the information, cf. below and the paper.) However, this bargaining with him or herself in a rational manner does not quite fit my understanding about how this kind of torture works. In particular, for the innocent X who happened to be detained, perhaps because another innocent Y under torture was forced to denounce others and happened to mention X's name. For the innocent X, and probably for the majority of torture victims according to the author's quoted literature, logic and bargaining are not relevant issues.

Contrary to the picture portrayed here, I believe that most often in intelligence interrogation the strategy is twofold: (1) to make the (guilty) victim give up values; and,

(2) to reduce his/her discernment by way of applying a battery of psychological and physical torture methods simultaneously and successively (well aware that from the outset all innocent people cannot be selected out and hence, some will have to be tortured).

The Inter-American Convention to Prevent and Punish Torture states, *inter alia*, in its definition of torture that: "Torture shall also be understood to be the use of methods upon a person intended to obliterate the personality of the victim or to diminish his physical or mental capacities, even if they do not cause physical pain or mental anguish." Obliteration of the personality surely means that the person loses their integrity, gives up their values and "betrays" their political ideals, their friends and fellows in the organisation that they belongs to. I don't see this as a matter of the victim bargaining with himself, but as a break-down.

Diminishing the mental capacity means the loss of discernment and not knowing exactly what is happening and what s/he is doing or saying, which is far from what in other contexts would be regarded as rational reasoning. Logic in the normal sense does not apply for the victim.

The aim of torture is that the victim loses all control at both levels.

In the author's model it is assumed that the interrogator's questions can be divided into two categories: leading questions to which the only possible answers are "yes" or "no" and "objective"/open-ended questions. The two types are kept as separate from each other in the model. It is further assumed that innocent detainees, i.e. those who have no information, cannot answer objective questions. I believe that the interrogation is different; the detainee is bombarded with both types of questions simultaneously making the distinction illusory. Moreover, unfortunately the innocent detainee can

in fact answer objective questions, e.g. give names of real persons who - falsely - are claimed to be implied in the crime under investigation, which reflects the break-down of the detainee, cf. above. But naturally, the innocent cannot provide useful intelligence information.

The conclusions of the study are clear. Some valid points are made about the dynamics of intelligence torture which are worth remembering; they are in accordance with common sense and with what we know and what is reported /quoted by the author: If torture is used to obtain "necessary information", e.g. to prevent terrorist acts or to fight organised crime, it has to be accepted that some persons with no affiliation to the criminals and having no relevant information are tortured since the interrogator can not know in advance which role and information the individual suspect has.

Likewise, the interrogator cannot know exactly when in the process of interrogation and torture the victim has given all the relevant information he has; hence, it is very likely that the interrogator continues the torture after having obtained all of the detainee's information. Moreover, at that point the interrogator may very well increase the intensity of torture because of the logic of torture: information will be obtained by way of torture, hence, if (additional) information is not obtained more torture must be applied.

Furthermore, the interrogator is faced with the difficulty of assessing the validity of the information obtained, having to sort out the valid from the fabricated, misleading and useless.

Accepting torture under certain circumstances will lead down a slippery slope and imply that a lot of innocent people will be tortured. Frequency and brutality of torture cannot be kept under control.

All of this goes against the arguments and assumptions of proponents of torture that reliable information can be obtained by way of torturing without having to torture innocent persons and without exceeding controlled limits, i.e. torture will not be more brutal than strictly necessary and will end when the information is obtained.

Schiemann's model has many similarities with the "ticking bomb scenario" where the preconditions, however, are extremely simplistic: (1) we know that a big terrorist attack will take place; (2) we have detained a person and we know that he is one of the perpetrators; (3) we know that he has the one piece of information necessary to prevent the attack; (4) we know that he will only give us the information if we torture him; and, (5) once having obtained the information we will be able to recognise it as the key we need to prevent the attack and torture will stop. The model described here is much more realistic: The torturer does not know whether the detainee is innocent or has the information required to prevent the crime; the torturer will not know if obtained information is the key to the resolution of the problem.

When comparing the analysis of the ticking bomb scenario with this scenario based on game theory, the only consideration that was new to me was that it is likely that the detainee will be tortured even harder after having given all the information he has because the torturer is not able to immediately realise that he has got all of the victim's information. However, a somewhat similar reasoning exists in the ticking bomb scenario where it is inferred that the victims who have no information to give are likely to suffer most torture. Hence, I am not quite convinced that the application of game theory in the torture scenario has provided substantial new information.

The well-known arguments against torture suffice: Torture is a crime under international law. It is very harmful to the victims, to democracy and to the torturers. But it is reassuring that the model, just like the analysis of the ticking bomb scenario, comes out with results that fit to our common arguments.

Response by the authors

John W. Schiemann PhD*

I appreciate the careful reading of my paper by all three commenters. I recognize that my paper is unusual for the *Torture Journal* in both argument and method. As a result, I am all the more grateful that practitioners from very different fields took the time to work through my paper. All three make good points and I welcome the opportunity to respond, even if only very briefly.

All the reviewers agree—and I with them—that torture is morally repugnant, illegal and should be prevented irrespective of efficacy. Thus, while we might all agree with Hans Petersen that "[t]he well-known arguments against torture" *should* "suffice," we also know that unfortunately they have not. Hence my attempt to, following the Alisons, "fully consider the counter arguments ... put forward by torture proponents" and confront them directly using their own logic. Of course, this approach is not without its own problems and limitations, and I respond to the three sets of comments in turn.

*) Department of Social Sciences & History, Fairleigh Dickinson University, Madison, New Jersey. Correspondence to: jws@fd.edu

Response to the Alisons.

The three questions raised by the Alisons are often asked and so good ones to raise here. I don't disagree with their answers, but would add some further considerations to the first two. With respect to why torture continues to be used despite its ineffectiveness, there is an additional explanation beyond retribution: what I called in my book the dangerous seduction of intuition. We have all felt pain and so we can all imagine being subjected to it and deciding to give up information to avoid (more of) it. The trouble is that intuition sometimes leads us astray and there is both observational and experimental evidence that people are able to withstand more pain than they anticipate *ex ante*.

As to the argument that torture is only supposed to "soften up" detainees so that rapport will elicit information later, the points made by the Alisons' are more or less those I make in my article and my book and so I agree. What I would point out in addition is that the Mitchell/Jessen program was pure theory. The Senate torture report makes it very clear that the torture as actually practiced differed little from its historical predecessors, with information sought at the time of torturing. Finally, the Alisons are, in my mind, spot on with their two counterarguments to the ticking bomb justification.

Response to Carle.

One dimension not captured by the model, Carle points out, is the Kafkaesque nature of the decision to use torture in the first place and the subsequent bureaucratic inertia to continue using it. Moreover, I, like Carle, noted that the Senate report revealed how much resistance to the program there was within the CIA and how many sincere but perhaps originally reluctant believers had been hoodwinked by others within the

CIA. Together they constitute a critical, but untold story deserving of further research.

It remains, however, very gratifying that the empirical experience of a real interrogator supports the more abstract model in multiple dimensions, from the strategic dynamic between detainee and interrogator to the the assumption by headquarters that a failure to provide information means the detainee is hiding information. My model gave the proponents every benefit of the doubt; Carle's testimony shows that gift is far too generous and so the worst outcomes are all the more likely.

Response to Petersen.

I believe there may be some misunderstanding of the model and its purpose insofar as the torture of an innocent does not conflict with the concept of a threshold because there are different thresholds for cooperative and innocent detainees in objective questioning. I hasten to acknowledge, however, that the responsibility for any such misunderstanding rests with me and my failure to clearly explain the model.

I appreciate the two conceptual points about the reality of torture conflicting with the model's portrayal of the victim's running inner dialogue as well as the distinction between objective and leading questioning. Even so I believe there are both broader and narrow caveats relevant here. First and more broadly, it is important to bear in mind both the purpose and nature of a model. Models simplify reality in order to identify and explicate important mechanisms or processes. Just as a wind-tunnel model of an airplane does not account for distasteful food, but attempts to model accurately flight behaviour, my model fails to capture every reality about interrogational torture in order to assess the proponents' argument on its

own terms. Proponents believe that a torture victim will eventually “decide” to give up information under threat of more pain and misery and so the model reflects that; it is not meant, nor does it claim, to be an accurate phenomenological representation of what it means to be tortured.

Second and more narrowly, my reading of accounts and memoirs of torture and the realistic but variegated outcomes of my model together suggest that the model’s simplifications are reasonable. Some torture victims actually do seem to go through something resembling that inner dialogue. The distinction between objective and leading questioning is not illusory—otherwise there would not have been rules and regulations on leading questions littered throughout the history of torture. Nor is it unhelpful insofar as the distinction helps bring out how easy it is to get bad information from leading questions and how innocents will be tortured under objective questioning. Indeed, Glenn Carle’s response makes clear that the assumptions of the model, however spare, are reasonable and accurate and tracing out their consequences is valuable even if it fails to capture everything about interrogational torture, let alone torture generally.

Torture was gradually excised from Continental European judicial systems from the sixteenth to the eighteenth centuries as a result of both arguments about human rights from Enlightenment *philosophes* as well as doubts about torture’s efficacy and necessity within a changing system of legal proof. Today, combining these efforts in rigorous arguments against torture’s efficacy may be the best way to push torture back into the dark recesses of history—where it belongs.