

# New Light on the Deflategate Controversy: Critical Technical Errors

Stephen McIntyre,  
Climate Audit,  
Toronto, Ontario  
May 31, 2016  
[stephengmcintyre@gmail.com](mailto:stephengmcintyre@gmail.com)

Submitted to Journal of Sports Analytics

## Abstract

There have been many critiques of the Wells Report on the Deflategate controversy, but all prior criticism (including my own) missed a simple but fundamental error in Exponent's technical report. When this error is corrected, there is sufficient data to preclude tampering after measurement of the footballs.

The newly identified error pertains to Exponent's simulations of Patriot ball preparation, an issue identified by Bill Belichick in his first press conference, but dismissed by Exponent and the Wells Report. In Exponent's simulations, they set football pressure to 12.5 psi before gloving, whereas the Wells Report reported that Patriot equipment manager Jastremski set pressure to 12.6 psi **after** gloving. Correctly simulated, Jastremski's method resulted in pressure at room temperature about 12.1-12.2 psi. This fully resolves the "unexplained" deflation.

Referee Anderson had two gauges, one of which (the Logo Gauge) read 0.38 psi too high. The only way in which referee Anderson could have missed the under-inflation was through using the Logo Gauge to measure Patriot balls. Exponent previously showed that Anderson had used the Non-Logo Gauge for Colt balls, implying that he inattentively switched gauges, identical to NFL officials at half-time.

Exponent also made an important error in their simulations of half-time transients using the Logo Gauge. When the transients are corrected, there is plenty of time at the beginning of the intermission during which Patriot observations and modeled transients are consistent, as well as later in the intermission for Colt balls.

The Deflategate controversy originated in scientific and technical errors. Appeal courts are poorly suited to resolve such errors. There is another way to resolve the controversy. The scientific community takes considerable pride in the concept of science being "self-correcting". When a scientist has inadvertently made an error, the most honorable and effective method of correcting the scientific record is to issue a corrected report, and, if such is not possible, retraction. Accordingly, even at this late stage, Exponent and/or Marlow should man up, acknowledge the errors and either re-issue corrected reports or retract. If either Exponent or Marlow do so, it is hard to envisage the Deflategate case continuing much further.

## 1 Introduction

In this article, I will first provide background as follows: a brief timeline of events from the original controversy to the legal appeals; a summary quantification of the amount of "unexplained" deflation actually in dispute; and a review of the battleground gauge controversy, including the NFL's determination that referee Anderson used the Non-Logo Gauge to measure Patriot balls.

I will then present the main results of the article: first, a re-statement of Figure 16 of the Exponent Report, showing the effect of correctly simulating Patriot ball preparation protocol; second, a re-statement of Figures 25 and 27 of the Exponent Report, culminating in a figure combining measurement of Patriot balls with the Logo Gauge and measurement of Colt balls with the Non-Logo Gauge. From these figures, many results follow, including the most fundamental finding that it can be shown that it is "very unlikely" that there was any tampering with the Patriot footballs after inspection by the referee.

I will then discuss several issues arising from the analysis, including whether there is any culpability for Patriots and/or Brady under the explanation presented here; how the scientific and technical consultants can fix the controversy arising from their errors; and some reflection on how “science says” rhetoric can be misused.

## 2 Background

### 2.1 Timeline

The New England Patriots played the Indianapolis Colts in the AFC conference final on January 18, 2015, won by the Patriots by a score of 45–7 (a fact oddly absent from the Wells Report). The game was played in the evening in cold rainy conditions with temperature at half-time falling to 48 deg F.

During the first half, the opposing team, the Indianapolis Colts, complained that Patriot footballs were under-inflated. At half-time, pressures of 11 Patriot footballs found to be under the 12.5 psi minimum prescribed by NFL rules. At the time, NFL officials were unaware that most, if not all, of the observed deflation was due to the effect of change in temperature on pressure according to the Ideal Gas Law.

On January 23, 2015, the NFL retained Paul, Weiss, a New York legal firm, to report on “the circumstances surrounding the use by the Patriots of footballs inflated at below-regulation air pressure levels during the AFC Championship Game, including whether Patriots personnel were involved in deliberate efforts to circumvent the Playing Rules”. Ted Wells led the lawyers from Paul, Weiss. Shortly afterwards, Paul, Weiss retained Exponent, a consulting engineering firm, “to provide scientific and analytical support for its investigation and help determine, based on the available data, whether it is likely that there had or had not been tampering with the Patriots footballs.” Paul, Weiss also retained Daniel Marlow of the Princeton University Department of Physics as a “special scientific consultant”.

On January 24, Patriot head coach Bill Belichick held a press conference in which he explained that footballs lost approximately 1 psi pressure due simply to the change to outdoor temperature. He also speculated that the ball preparation technique used by the Patriots for the AFC Championship Game, heavy manual gloving, could also have contributed to the observed pressure drop. Numerous contemporary articles also pointed out that football pressures were subject to the Ideal Gas Law (e.g. Glanz, 2015) and that simple application of the Ideal Gas Law to the change in temperature between the locker room and the field would explain all or most of the pressure drop. The point was recently re-iterated in the *Professors’ Amicus Brief (2016)* in connection with the most recent appeal.

On May 6, 2015, Paul, Weiss released their “Investigative Report Concerning Footballs Used During the AFC Championship Game on January 18, 2015” (the Wells Report), including, as an exhibit, Exponent’s report of the same date entitled “The Effect of Various Environmental and Physical Factors on the Measured Internal Pressure of NFL Footballs” (the Exponent Report). The Exponent Report provided the first information on actual pressures, which showed less deflation than the original leaked information (which the NFL left uncorrected.) It also disclosed the surprising information that referee Anderson had two different gauges (the Logo and Non-Logo Gauge), with one of the gauges (Logo) being biased approximately 0.38 psi too

high. Subsequently, Anderson's pre-game gauge selection would soon become a major issue. Exponent also reported for the first time that, despite the heightened scrutiny in the half-time measurements, NFL officials had inadvertently switched Logo and Non-Logo gauges between Patriot and Colt balls – the bias between the two gauges only later being recognized.

Exponent carried out a wide variety of tests and simulations on the footballs. Although many critics have accused the NFL of ignoring the Ideal Gas Law, Exponent discussed its effect at length, even arriving at a similar estimate of its effect (~1 psi) as Belichick had reported at his original press conference. Exponent included a (purported) simulation of Patriot ball preparation technique, concluding that it resulted in a temporary increase of pressure of ~0.7 psi, but that this effect wore off after about 15-20 minutes. From this test, Exponent rejected Belichick's suggestion that Patriot ball preparation might have contributed to the observed deflation. Exponent also simulated footballs under conditions of extreme wetness, finding that pressure of wet footballs was about 0.3-0.5 psi lower than corresponding dry footballs.<sup>1</sup> Exponent examined many "exemplar" gauges, none of which had a similar bias to the Logo Gauge. From this analysis (described in more detail below), they concluded that referee Anderson had used the Non-Logo Gauge for pre-game measurements. Nonetheless, they carried out simulations of changes in pressure for the change in temperature from indoors to outdoors (71 deg F to 48 deg F) and as footballs warmed up at half-time (~73 deg F) for both Logo and Non-Logo gauges. The half-time pressure transients from these simulations showed an increase of pressure in a close-to negative exponential function, increasing by about 0.067 psi/minute in the first seven minutes and by about 0.032 psi in the last seven minutes.

From their analyses, Exponent stated:

the data did not provide a basis for us to determine with absolute certainty whether there was or was not tampering as the analysis of such data ultimately is dependent upon assumptions and information that is not certain. However, based on all of the information provided to us, particularly regarding the timing and sequencing of the measurements conducted by the game officials at halftime, and on our testing and analyses, we conclude that within the range of game characteristics most likely to have occurred on Game Day, **we have identified no set of credible environmental or physical factors that completely accounts for the additional loss in air pressure exhibited by the Patriots game balls as compared to the loss in air pressure exhibited by the Colts game balls measured during halftime of the AFC Championship Game.**

Relying on the Exponent Report, the Wells Report endorsed the above findings, which they characterized as "tend[ing] to support a finding that human intervention may account for the additional loss of pressure exhibited by the Patriots balls":

Based on the testing and analysis, however, Exponent concluded that, within the range of likely game conditions and circumstances studied, they could identify no set of credible environmental or physical factors that completely accounts for the Patriots halftime measurements or for the additional loss in air pressure exhibited by the Patriots game balls, as compared to the loss in air pressure exhibited by the Colts game balls. Dr. Marlow agreed with this and all of Exponent's conclusions. This absence of a credible scientific explanation for the Patriots halftime measurements **tends to support a finding that human intervention may account for the additional loss of pressure exhibited by the Patriots balls.**

On May 12, 2015, citing the Wells Report, the NFL found against Brady and suspended him for four games. Tom Brady and the NFL Players Association immediately appealed.

---

<sup>1</sup> Exponent did not give a specific number, but these values can be estimated from Figures 25 and 27.

A hearing was held on June 23, 2015, though the transcript did not become available until September 2015, when it was released in connection with subsequent legal proceedings. The following technical consultants were examined during the hearing: for Brady, Edward Snyder of Yale University; for the NFL, Robert Caligiuri and Duane Steffey of Exponent and Daniel Marlow of Princeton University (Department of Physics). The hearing was lengthy, but not without its droll moments, of which Goodell's examination of Brady in connection with the 2006 rule change on ball preparation stands alone. Goodell wanted to get to the bottom of the 2006 rule change:

Q. And did the Competition Committee adopt this rule change to let the quarterbacks prepare their balls?

A. Yes.

COMMISSIONER GOODELL: Mr. Kessler, who did this come from?

MR. KESSLER: The petition?

COMMISSIONER GOODELL: Yes.

MR. KESSLER: The quarterbacks and the League all signed it.

COMMISSIONER GOODELL: I know. I signed it. But who did it come from?

Brady tried to explain, but his answer merely added to Goodell's puzzlement:

THE WITNESS: Peyton.

Brady's answer was too allusive for Goodell, who asked for more information on the identity of the mysterious originator of the rule change:

COMMISSIONER GOODELL: Peyton?

Fortunately, Brady was able to complete the identification for the perplexed commissioner:

THE WITNESS: Manning.

On July 23, 2015, the NFL issued its "Final Decision on Article 46 Appeal of Tom Brady", upholding the suspension. Commissioner Goodell ratcheted up the findings of the Wells Report, stating that "at least a substantial part of the decline was the result of tampering", a statement appearing nowhere in the Wells Report.

the full extent of the decline in pressure cannot be explained by environmental, physical or other natural factors. Instead, **at least a substantial part of the decline was the result of tampering.**

Goodell's decision specifically addressed the controversy over which gauge had been used to measure Patriot footballs pre-game.

There was argument at the hearing about which of two pressure gauges Mr. Anderson used to measure the pressure in the game balls prior to the game. The NFLPA contended, and Dean Snyder opined, that Mr. Anderson had used the so-called logo gauge. On this issue, I find unassailable the logic of the Wells Report and Mr. Wells' testimony that the non-logo gauge was used because otherwise neither the Colts' balls nor the Patriots' balls, when tested by Mr. Anderson prior to the game, would have measured consistently with the pressures at which each team had set their footballs prior to delivery to the game officials, 13 and 12.5 psi, respectively. Mr Wells's testimony was confirmed by that of Dr. Caligiuri and Professor Marlow. As Professor Marlow testified, "There's ample evidence that the non-logo gauge was used".

Goodell also rejected AEI's suggestion that referee Anderson had switched gauges between measurement of Colt and Patriot balls (as NFL officials had inadvertently done in their half-time measurements).

The NFL then filed a motion in New York for a declaratory judgement, while the NFLPA appealed in Minnesota, with the New York proceeding going forward. On September 3, 2015, Judge Berman of the 2<sup>nd</sup> Circuit reversed Goodell's suspension, permitting Brady to play the full 2015 season.

The NFL appealed. In this appeal, the scientific issues were not addressed; it instead focused exclusively on Commissioner Goodell's powers under the collective bargaining agreement. Professor Blecker filed an amicus brief, in which he drew attention to the scientific issues. On April 24, 2016, the 2<sup>nd</sup> Circuit Court of Appeal reversed J Berman's decision in a 2-1 decision, re-instating Brady's suspension.

Brady and the NFLPA immediately appealed, seeking a re-hearing en banc by the 2<sup>nd</sup> Circuit Court of Appeal. Amicus briefs touching on scientific issues have been filed by the New England Patriots and by a group of professors of engineering and physics (the *Professors' Amicus Brief*).

## 2.2 Quantification of the Deflation in Dispute

Exponent itemized the main factors affecting pressure drop as follows:

- Deflation arising from cold on-field conditions (the Ideal Gas Law)
- Differential warming of dry balls due to Colt balls being measured later in the half-time interval
- Additional deflation in wet balls. Exponent determined that the volume of the footballs did not change materially, so the effect is necessarily due to lower internal temperatures in wet footballs;

As others have pointed out (e.g. *Professors' Amicus Brief, 2016*), uncertainty is attached to each of these effects, with the amount of uncertainty varying by effect. Unfortunately, the Exponent Report failed to provide a quantitative summarization of the relative size of the various effects. This omission has permitted the flourishing of various misinterpretations of the amount of unexplained deflation. For reference, I've attempted to extract central values for each effect for each team as estimated in (or consistent with) the Exponent Report and then provide a reconciliation to quantify the amount of "unexplained" deflation (see Figure 1 below).

Based on observed half-time pressures from the unbiased Non-Logo Gauge, the pressure drop for Patriot balls from pre-game values measured by referee Anderson was about 1.39 psi, about 0.71 psi more than the corresponding Colt pressure drop (0.68 psi).

The estimated pressure drop for dry footballs under the Ideal Gas Law is uncontroversial and well-constrained. For a drop in temperature from 71 deg F to 48 deg F, the expected pressure drop was 1.08 psi for Patriot balls and 1.10 psi for Colt balls, with the slight difference arising from the slightly higher inflation of Colt balls to 13 psi. While the NFL was clearly unaware of the Ideal Gas Law when it commenced half-time measurements, Exponent was well aware of the effect and fully accounted for it.

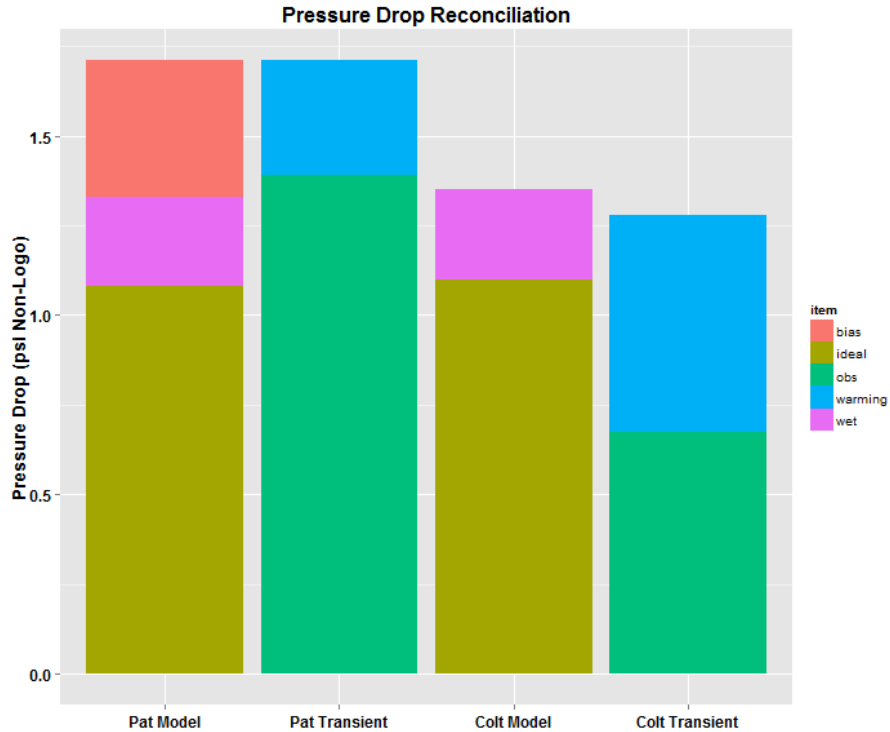


Figure 1. Reconciliation of Patriot and Colt pressure drops. In the right column of each pair, estimated warming through the intermission is added to the observed pressure drop to estimate the pressure drop at the start of the half-time intermission. In the left column of each pair are shown the pressure drop for dry balls (limegreen), an estimated average additional drop for wet balls and, for Patriot balls, the additional deflation arising from re-setting pressure after gloving.

More uncertainty attaches to the pressure gain during the half-time intermission. In Figures 25 and 27, Exponent showed four pairs of transients, with each pair consisting of transients for dry and extremely wet footballs. While there is some variability in the coefficients for dry balls, the main uncertainty attaches to the average time of measurement, since NFL officials did not record the measurement times. It is known that Patriot balls were measured first. Exponent attempted to replicate the Game Day sequence. A plausible average measurement time of ~4 minutes (the central value of Exponent Figure 30) results in ~0.32 psi warming. An uncertainty range of  $\pm 90$  seconds is equivalent to ~0.095 psi. Colt footballs were measured later in the intermission, but there is a bizarre uncertainty about whether they were measured immediately after Patriot balls (and before reflation of Patriot balls) or whether they were measured at the end of the intermission (after reflation.) In Figure 1, following the central value of Exponent Figure 30, an average measurement time of ~8.5 minutes (following) results in ~0.60 psi warming. If Colt balls were measured toward the end of the intermission (as plausibly suggested by numerous commentators on the grounds that Colt measurements were interrupted after only four balls by the end of the intermission), there would have been ~0.72 psi warming, a difference of ~0.12 psi. In either case, the discrepancy between Patriot and Colt pressure drops is reduced to ~0.3-0.4 psi.

The third major factor was wetness. Exponent's simulations conclusively showed that wet footballs experienced additional pressure loss. In other (uncontroversial) simulations,

Exponent showed that there was no relevant change in volume. The Ideal Gas Law therefore implies that the internal ball temperature of the wet footballs was lower than that of dry footballs, an effect consistent with the well-known phenomenon of evaporative cooling. In the absence of data on the relative wetness of Patriot and Colt footballs, Exponent elected to simulate two extremes: dry and extremely wet footballs, with their wetting protocol attempting to be even more extreme than game conditions, showing an envelope between the two extremes. Uncertainty associated with relative wetness has been a source of controversy, but, in my opinion, Exponent's simulations reasonably constrain the uncertainty of the average. From the transients depicted in their Figures 25 and 27, wet-dry differentials from each pair can be calculated. While there is considerably more variability in the shape of the wet-dry differential transients than one would like, Exponent's results show a maximum differential of ~0.5 psi. For first-order reconciliation, I've assumed 50% of the maximum value for both Patriot and Colt balls, each value obviously having uncertainty, but the uncertainty appears to me to be in the order of 0.1-0.2 psi, rather than something much higher. Patriot defenders have pointed out that the Patriots had more ball possession in the first half, including a long closing drive, and therefore could have been wetter.

The first-order reconciliation on these assumptions is shown in Figure 1, with Patriot data shown in the first two columns and Colt data in the third and fourth columns. In the right column of each pair, I've shown the total of the observed deflation (Non-Logo Gauge) and estimated half-time warming as an estimate of deflation at the start of the half-time intermission. In the left column of each pair, I've shown estimates of the physical contributions from Ideal Gas Law cooling to 48 deg F and wetting (as discussed above). For Colt balls, the first-order values reconcile reasonably closely without tuning. Plausible combinations of later measurement time and/or slightly less wetting would reconcile completely.

For Patriot balls, the same factors yield a discrepancy of ~0.37 psi. In the first column, I've also illustrated the bias of the Logo Gauge, showing that the amount of unexplained Patriot deflation is almost exactly equal to the bias of the Logo Gauge. Due to Exponent's failure to clearly reconcile the various deflation contributions, this *coincidence* is not evident in the Exponent Report, but it is the sort of thing that ought to trouble any analyst. In the absence of any clear rationale for a competitive advantage arising from ~0.37 deflation, it's also a coincidence that ought to have troubled Exponent, Wells and Goodell.

In any event, the amount is indisputably a second-order effect, with deflation from the Ideal Gas Law and wetness amounting for nearly 80% of the observed deflation. Goodell's characterization of this increment as "at least a substantial part of the decline" was an unjustified embellishment of findings of Exponent and Wells.

### 2.3 *The Gauge Controversy*

Almost immediately after publication of the Wells Report, MacKinnon (2015) observed that, if referee Anderson had measured Patriot footballs pre-game with the Logo Gauge, Patriot ball deflation could be "explained on the basis of physical law, without manipulation". Hassett et al. (2015) arrived at the same conclusion, additionally observing that it was possible that referee Anderson had not used the same gauge to measure both Colt and Patriot balls, in attentively switching gauges as NFL officials had done at half-time, with the scenario of interest



being use of the Logo Gauge to measure Patriot balls and Non-Logo Gauge to measure Colt balls. Similar points have been made in subsequent critiques (Blecker Amicus Brief, 2015; Professors' Amicus Brief, 2016).

Exponent was fully aware of this line of argument and a considerable portion of the Exponent Report was clearly written in anticipation of such criticisms. Their main rebuttal came from their finding that Anderson had measured balls pre-game with the Non-Logo Gauge, a finding that mooted the apparent reconciliation using the Logo Gauge. Exponent was aware of and reported that Anderson recollected that he had used the Logo Gauge (allowing that his recollection could be faulty), but rejected Anderson's recollection based on their analysis.

Most critics of the Wells Report have ignored this topic, a lacuna which leaves the Wells Report unchallenged on an essential topic. The few critics of the Wells Report who have challenged Exponent's reasoning on these issues (notably Snyder at the Appeal Hearing and the Blecker Amicus Brief) resorted to rather unsatisfying speculation. Under the analysis presented in this articles, their speculative explanations can be shown to be incorrect, though their search for an explanation was correct.

### 2.3.1 *Exponent's Argument*

Exponent's analysis was grounded on the observation that the pressure settings recalled by both Patriot and Colt personnel were similar to the pressures recalled by referee Anderson:

According to information provided by Paul, Weiss, personnel from both the Patriots and the Colts recall gauging the footballs for their teams to pressures at or near 12.5 psig and 13.0 psig, respectively, prior to providing the balls to Walt Anderson. Each team used its own gauge to adjust the final pressures before presenting the balls to the referee, who used a gauge different from either used by the two teams to measure the pressure in the footballs. Walt Anderson recalled that according to the gauge he used (which is either the Logo or Non-Logo Gauge), all of the Patriots and Colts footballs measured at or near 12.5 psig and 13.0 psig, respectively, when he first tested them (with two Patriots balls slightly below 12.5 psig). This means that the gauges used by the Patriots and the Colts each read similarly to the gauge used by Walt Anderson during his pregame inspection.

Even Blecker agreed with this logic thus far:

Exponent's logic: Since the Referee's gauge essentially confirmed the accuracy of each team's pre-set ball pressure, we can assume he in fact used whichever of his two gauges more nearly matched the teams' own gauges. That makes sense.

During its research, Exponent had analyzed dozens of new gauges ("exemplar gauges") that it purchased or was provided. All were relatively accurate; none of them had similar bias to the Logo Gauge which read consistently ~0.38 psi too high. Exponent argued that it was very unlikely that both Patriots and Colts could have used gauges that were similarly out of compliance to the Logo Gauge, from which they deduced that Anderson must have used the Non-Logo Gauge:

As was shown earlier, the Logo Gauge typically read ~0.3–0.4 psig higher than the Non-Logo Gauge on Game Day, according to the data recorded at halftime. It was shown by our experiments that the Non-Logo Gauge was relatively accurate in an absolute sense when compared after the fact against the known calibration of the Master Gauge. ...

It has been shown that the Logo Gauge consistently reads higher than all other gauges analyzed in this investigation. As a result, it is very unlikely that the Logo Gauge would have read similarly to the gauges used by each team. **Therefore, it is most likely that the gauge used by Walt Anderson prior to the**

**game was the Non-Logo Gauge, which read similarly to the Master Gauge and other gauges tested during the investigation.**

### 2.3.2 *Wells Report*

The Wells Report (WR, 116) cited Exponent’s finding about the gauges as follows:

Exponent relied upon this information, as well as the fact that during the testing the Non-Logo Gauge never produced a reading higher than the Logo Gauge, to conclude that Walt Anderson most likely used the Non-Logo Gauge to inspect the game balls prior to the game,

It additionally endorsed Exponent’s findings on this point as showing that Anderson’s use of the Non-Logo Gauge was “more probable” and “most likely”:

For the reasons described in Section VII.B, we believe it is more probable that Anderson used the Non-Logo Gauge for his pre-game measurements. (WR, 52 n30)

As noted above, we also believe that Walt Anderson most likely used the Non-Logo Gauge prior to the game. (WR, 67 n37)

### 2.3.3 *Snyder’s Attempt to Rebut Exponent*

At the Appeal Hearing, Edward Snyder, the statistical consultant to Brady and the NFLPA, proposed that the Patriots could have used an older gauge with a similar bias to Anderson’s Logo Gauge. Under this scenario, Patriots inadvertently tendered balls to Anderson that were under-inflated under a true gauge, but the under-inflation was missed by Anderson who failed to notice because he was using a gauge that read too high, failed to notice. In this scenario, the balls were slightly under-inflated at pre-game measurement, but no tampering after measurement. While Snyder misdiagnosed the reason for under-inflation when the balls were tendered, other elements of the reasoning hold up.

Snyder presented this as follows (*Hearing*, 217:12):

Q.... And my question is, do you understand that the Patriots say that the psi level that they set the balls to before the game was 12.5 or 12.6? Is that your understanding of the Patriots' position?

A. With their -- with whatever gauge they used, that's their understanding, you are right.

Q. So the answer to my question is yes, that's your understanding?

A. Yes, but we don't know what that gauge was. We don't know if that was giving accurate measures or not. It means that if they used a gauge that was like the logo gauge, they would have delivered balls that were 12.5 and they didn't know it.

NFL lawyer Lorin Reisner confronted Snyder with Exponent’s survey of “dozens” of gauges, which had not yielded any gauges with bias comparable to the Logo Gauge. Snyder argued that Exponent’s survey had been limited to new gauges more or less identical to the Non-Logo Gauge and that Exponent’s failure to study the properties of older gauges left open the possibility that other older gauges might have gone off specification commensurate with the Logo Gauge (*Hearing*, 219:2):

But we don't know what Patriots' gauge was used and there is no basis in the report... You don't have the Patriots' gauge. You don't know if that was an older gauge. There's evidence to indicate that older gauges read higher than new gauges. Exponent collected a bunch of new gauges and said, oh, new gauges, fine. No surprise there. I'm not -- I'm not quarreling with this. It's just the major point here is there are just so many uncertainties.

### 2.3.4 Wells and Caligiuri Respond at the Appeal Hearing

At the Appeal Hearing, Wells testified on the Logo/Non-Logo issue at far greater length and with much greater vehemence than in the Wells Report itself, with his evidence on the point occupying no less than 10 pages of the transcript. (Hearing, 290-302).

Wells emphasized that he had made an “express finding” that Anderson had used the Non-Logo Gauge and that this was not “uncertain” (Hearing, 290:14):

Q. And Mr. Wells, another point you came to as being uncertainty, you wrote was whether or not when the initial testing was done before the game, whether that was done by the logo gauge or the non-logo gauge, right? You concluded that was uncertain?

A. No, no. I made an express finding and so did the -- the experts made a finding and I -- and when I say "I," I mean collective "I," my team, we made an express finding that the non-logo gauge is the gauge that was used by Walt Anderson when he tested the balls. That is an express finding in the report.

Wells re-iterated Exponent’s argument, characterizing Patriot information on how they set pressure as follows (Hearing, 293:3), but failing to mention that Patriots set pressure to 12.6 psi **after** gloving:

Now, the evidence we have is that the Patriots were emphatic with us that they set their balls at 12.5 or 12.6. That testimony came from Mr. Jastremski and it also came from Mr. Brady. Our balls are coming in at 12.5 or 12.6. So that's the Patriots. So I assume for the AFC Championship Game, the Patriots are set. They know where they are setting their balls. They have told me they are 12.5, 12.6.

Wells later (Hearing, 294:5) expressed the other leg of Exponent’s argument as follows:

Now, how do you get to what gauge he used? The only way Walt Anderson could get to 12.5 for the Patriots and 13 for the Colts is if he used the non-logo gauge. And that is because the logo gauge always reads .3 to .4 higher. It is consistent.

Wells colorfully expressed the practical improbability of both Patriot and Colt gauges being “out of whack like the logo gauge” as a “lightning strike” (Hearing, 295:2):

And the only way you could get those measurements where Walt says he saw just what the Patriots saw and what the Colts saw is with the non-logo gauge. And that's why we made that finding. Now, maybe lightning could strike and both the Colts and the Patriots also had a gauge that just happened to be out of whack like the logo gauge. I rejected that.

Kessler pointed out to Wells that they had never found either the Patriot gauge or Colt gauge and therefore could not verify for sure that either of them was similarly out of compliance. Because Exponent’s exemplar gauges had all been new, Kessler raised the possibility that the reading of an older gauge might have become biased similarly to the Logo Gauge (Hearing, 296:11).

Q. If their gauges read like the logo gauges because they were older gauges that were given by Wilson and may have looked just like the logo gauge, then they might read like the logo gauge if that was true?

Wells rejected this hypothesis as improbable, again using his “lightning strike” metaphor:

A. That's what I mean if lightning were to strike and what you would have to have happen in terms of my analysis, you would have to have had both teams for that Championship Game had gauges that were .3 to .4 off and then that all flowed into Walt Anderson using the logo gauge which was .3 to .4 off. And I don't think that happened and that's what I ruled. I think what I ruled is totally -- not only do I think it's correct, I think it's reasonable.

Kessler challenged Wells with referee Anderson's recollection of having used the Logo Gauge. Wells testified that he found that Anderson's recollection was mistaken because of the "scientists" and their "scientific tests":

So I say Walt Anderson says it is his best recollection that he used the logo gauge. We then did tests that showed that there is consistent uptick on the logo gauge of .3 to .4. **The scientists, the Exponent people say they believe based on their scientific tests that the non-logo gauge was used.** I have a ruling that says there's uncertainty, but I am making a ruling as a finder of fact, because that's my job as the judge, that it's more probable than not that the non-logo gauge was used by Walt Anderson. That is set forth in those words or substance in both my report and in the Exponent report.

Later in the hearing, Kessler challenged Exponent's Caligiuri with the possibility that the Patriot gauge was an older gauge that had got out of calibration similarly to the Logo Gauge. Caligiuri curtly rejected the probability of this: (*Hearing*, 380:19).

Q. No, that over time they got out of calibration.

A. All the gauges got out of calibration by the same amount?

Q. Over the same period of time.

A. I would say that's pretty highly unlike.

Q. You think that's unlikely? Did you do any test for that?

A. How would we test that? We tested the logo gauge and found that it reads very repeatedly .3 to .45 above the master gauge. And we tested hundreds of gauges that we got, exemplar gauges. Yes, you are right, they are new. And they all read what the master gauge said they should be.

Caligiuri was also asked about the possibility of Anderson changing gauges between Colt and Patriot balls, a possibility raised by AEI which reduced the problem of both Patriots and Colts having gauges "out of whack like the logo gauge" to only the Patriots – a substantial reduction in improbability. Caligiuri vehemently contested even the possibility of, adding the astounding statement that Exponent had been "told to assume that did not happen":

there's never been any indication that Walt Anderson switched the gauges in the middle of his pre-game measurements. We had no indications. **We were actually told to assume that that did not happen** because there was no evidence that that happened outside of AEI. We take those two scenarios off the table. [my bold]

It is hard to contemplate a legitimate reason for excluding that scenario from the scope of technical analysis. The decision becomes even more questionable in retrospect, given the demonstration in this article that it is likely that such a switch took place. The New England Patriots' Amicus Brief challenged the unfairness of the NFL withholding documents giving instructions to Exponent. Had such documents been available, this issue might have attracted attention on a more timely basis.

### 2.3.5 *Subsequent Amicus Briefs*

Blecker's amicus brief discussed the gauge controversy. Blecker reminded readers that referee Anderson had recalled using the Logo Gauge and that Exponent and Wells had concluded that his recollection was mistaken. Blecker pointed out that it was "very odd" that the gauges of both teams had "disappeared" and were unavailable to confirm or rebut Anderson's memory. Blecker went on to satirize Exponent's collection of new gauges identical to the Non-Logo Gauge, without studying properties of older gauges such as the Logo Gauge. While Blecker's points are fair enough, these points had been previously raised at the Appeal Hearing.

Despite many sensible observations, unfortunately, the Professors' Amicus Brief did not squarely address the NFL's gauge findings. They characterized the NFL's position on the Non-Logo Gauge as an "assumption", which, if varied (together with a plausible temperature variation), led to a different result:

Likewise, in asserting that the increment could be as high as 0.53 psig, the league assumed that the referee used a particular gauge. JA147...If these two assumptions were changed to be consistent with the temperature assumption made elsewhere and with the information provided by the referee", the increment of pressure loss is as low as zero. ...Had the league made these two different assumptions alone, the results would have vindicated Mr Brady"

The "particular gauge" is, of course, the Non-Logo Gauge. However, rightly or wrongly, the NFL had not simply "assumed" that Anderson had used the Non-Logo Gauge. From their perspective, it was a "logical inference", resulting in a "finding" rather than a mere assumption. The same point had been previously made in similar form by MacKinnon, Hassett et al and Snyder and argued at length at the Appeal Hearing.

### 3 Analysis

I will now present the main results of the article: first, a re-statement of Figure 16 of the Exponent Report, showing the effect of correctly simulating Patriot ball preparation protocol; second, a re-statement of Figures 25 and 27 of the Exponent Report, culminating in a figure combining measurement of Patriot balls with the Logo Gauge and measurement of Colt balls with the Non-Logo Gauge. From these figures, many results follow, including the most fundamental finding that it can be shown that it is "very unlikely" that there was any tampering with the Patriot footballs after inspection by the referee.

#### 3.1 Issue 1. Simulation of Patriot Ball Preparation

The effect of the Patriot manual gloving technique was in Exponent's terms of reference and was discussed at length in the Exponent Report (Figure 16; pages 33ff), in the Wells Report (pp 11, 50, 112, 119, 120, 131) but only touched on tangentially in the Hearing.

##### 3.1.1 Jastremski's Procedure for AFC Game

Ball preparation for the AFC Championship Game was not done according to usual Patriot procedures. Brady testified that, because heavy rain was expected for the AFC Championship Game, he was concerned that their usual ball preparation with Lexol would result in slippery footballs and that he asked Jastremski to prepare the balls with heavy manual gloving, a procedure that Jastremski's predecessor had used several years earlier, but which had apparently not been used during Jastremski's tenure as equipment manager (*Hearing*, 69:3).

Jastremski's implementation of the procedure was described in the Wells Report (WR, 50) as follows:

He [Jastremski] and other members of the equipment staff then "gloved" the footballs, spending between 7 and 15 minutes vigorously rubbing each ball....Jastremski told us that **he set the pressure level to 12.6 psi after each ball was gloved** and then placed the ball on a trunk in the equipment room for Brady to review. [my bold]

While the Wells Report and Exponent Report typically track one another on technical details, the information that Jastremski set the pressure level to 12.6 psi "**after** each ball was

gloved” was not reported in the Exponent Report or implemented in their simulation of Patriot ball preparation.

Although the ball preparation for the AFC Championship departed markedly from their usual procedure and from procedures that had been used during Jastremski’s tenure, the Wells Report (WR, 37) implied that Jastremski’s process was standard practice (an assertion that is contradicted by other facts in the report):

The Patriots have developed a process for the preparation of game balls in accordance with the preferences of Tom Brady, who has been the team’s starting quarterback for over thirteen years.

### 3.1.2 Exponent Report Figure 16

Exponent carried out simulations on the potential impact of Patriot ball preparation (“vigorous rubbing”) (see Exponent, 33-34 and Figure 16, also shown below). In their text summarizing Figure 16, Exponent stated that the “pressure inside the football rises throughout the rubbing process”, that “the non-linear nature of the rise in pressure is presumably due to variations in the operator's rubbing throughout the period”, that “the maximum rise in pressure was approximately 0.7 psig, which occurred at the end of the rubbing process” and that “the pressure returns to its initial state approximately 30 minutes after the cessation of the rubbing.” All these features can be easily identified in Figure 16.

For this figure, Exponent had set ball pressures to 12.5 psi **before** the gloving procedure. This was confirmed by Exponent’s Caligiuri at the hearing. (*Hearing*, 397:16). Because the effect of the rubbing wore off in about 15-20 minutes well prior to measurements by referee Anderson, Exponent concluded that the “the vigorous rubbing described by Patriots personnel does not appear to have had an impact on the pressures measured in the Patriots footballs either prior to the game or at halftime”.

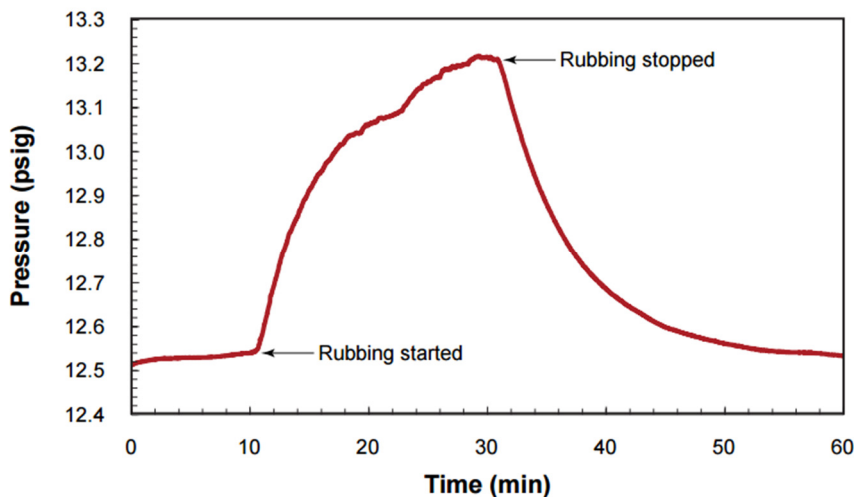


Figure 2. Exponent Figure 16. Original Caption: The pressure as a function of time while a football is being vigorously rubbed.

### 3.1.3 Re-Stated Figure 16

As noted above, Exponent's simulation omitted an essential feature of Jastremski's protocol: setting the pressures to 12.6 psi **after** gloving. In Figure 3, I've annotated Exponent's Figure 16 to show the difference.

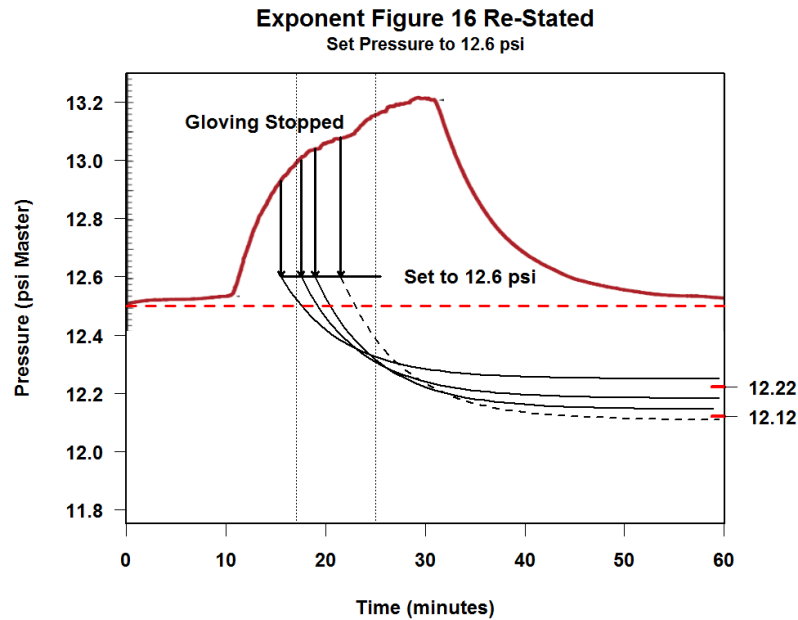


Figure 3. Re-statement of Exponent Figure 16. Red: Exponent transient shows effect of rubbing to increase pressure, together with a decline after rubbing stopped. Black: shows reduction in pressure from re-setting to 12.6 psi, followed by transients as ball temperature returns to room temperature. Twenty seconds allowed for setting gauge in above transients. Dotted vertical lines show 7-15 minutes from start of rubbing reported by Jastremski. Logo gauge values of 12.5 and 12.6 psi are shown on right axis, deducting the bias of ~0.38 psi.

Jastremski stated that he gloved the balls for 7 to 15 minutes, after which he set the pressure to 12.6 psi. During the gloving procedure, the pressures and temperatures would increase, more or less as depicted in Figure 16. Under the Ideal Gas Law, the observed pressure increase corresponds to an increase of internal ball temperatures to ~80-83 deg F. If the balls had a pressure of ~12.5 psi *before* rubbing, Jastremski's setting of the pressure of the *warm* (80-83 deg F) balls to 12.6 psi required deflation by approximately 0.35 psi. I've illustrated this step through downward arrows at 5, 7, 8.5 and 11 minutes after the start of gloving.

Following the pressure of the warm balls being set to 12.6 psi, the balls will experience a similar cooling transient to that depicted in Figure 16, but from a lower starting point. The re-setting of pressure would not be instantaneous with the termination of rubbing, but the interval of time is not likely to have been long either. I've allowed 20 seconds in the above transients. The warmer the ball from rubbing at the time of re-setting, the lower the pressure of the ball when it returns to ambient temperature.

On the right axis, I've marked pressures of 12.12 and 12.22 psi, which correspond to readings of 12.5 and 12.6 psi by the Logo Gauge (which read ~0.38 psi too high). To fall within

this interval, the footballs would have to have been set to 12.6 psi at a ball temperature between 78.5 and 80.5 deg F– a range of values that is consistent with Jastremski’s rubbing protocol.

### 3.1.4 Revised “Logical Inferences”

Exponent stated that there was enough information to make “logical inferences” on gauge usage – a point on which I agree. However, the above information shows how Exponent arrived at exactly the wrong inference in respect to Patriot footballs.

Assuming for now that Jastremski’s gauge was relatively unbiased (as argued by both Exponent and Wells, and, with which I agree), Jastremski’s protocol would result in Patriot footballs having (true) pressures of about 12.12-12.22 psi when measured by referee Anderson at room temperature –marked in red on right axis in Figure 2. If Anderson had measured these balls with the Non-Logo Gauge, he would have noticed their under-inflation. The *only* way in which Anderson could have failed to notice the under-inflation of Patriot balls prepared with Jastremski’s method was if he had measured them with the Logo Gauge. Pressures for Colt balls were set at room temperature and Exponent’s reasoning in respect to Colt balls, with which I agree is unaffected: i.e. that Anderson used the Non-Logo Gauge for Colt balls.

A corollary to the above inferences is that Anderson must have inattentively switched gauges between measurement of Patriot and Colt balls. The possibility (though not the fact) of such a switch is convincingly established by events at half-time, where, despite heightened scrutiny, Exponent concluded (reasonably) that NFL officials had inattentively switched gauges between measurements of Patriot and Colt balls. Indeed, if Anderson returned a gauge to his pocket after measurement of one set of footballs, the selection of gauge for the other set of footballs would be random, with switching of gauges having equal probability as using the same gauge. Because of this form of uncertainty, the determination of which gauge was used to measure Patriot balls has to be done on its own merits.

Correctly construed, using the same form of reasoning as originally used by Exponent (and endorsed so enthusiastically by Wells), the data shows that it is “very likely” that Anderson used the Logo Gauge to measure the Patriot footballs and the Non-Logo Gauge to measure Colt footballs.

### 3.1.5 A Near Miss on the Rubbing Effect

Although none of the parties at the Appeal Hearing appears to have fully understood the precise implication of the gloving protocol, Jeffrey Kessler, counsel to the NFLPA and Brady, asked the following exactly on point question to Exponent’s Robert Caligiuri (*Hearing*, 400:10), which, if answered correctly, might have unraveled the NFL argument:

**Q.** Okay. So if they set their psi at 12.5, after rubbing, that was their procedure, under your analysis, how much below 12.5 would it drop as the rubbing effect wore off?

The correct answer to this question was approximately 12.17 psi. Had Caligiuri given this correct answer, it is possible that the NFL’s case might have unraveled. However, Caligiuri’s answer was unresponsive, merely saying that the rubbing effect wore off in 20-30 minutes – a point not at issue in Kessler’s question.

The rubbing effect is worn off within about 20 or 30 minutes of when you started.

Goodell then interjected with an observation that had nothing to do with Kessler’s question:



COMMISSIONER GOODELL: Mr. Kessler, didn't Mr. Brady testify that he picked the ball up three to four hours before the game?

The discussion then went off on a tangent. While it is not clear that Kessler realized the importance of this question, by the time that discussion returned from the tangent (*Hearing*, page 402 line 8), Kessler had lost his train of thought and unwisely abandoned the topic, incorrectly observing that the facts were “too confused” and the topic not “significant to the overall analysis”.

MR. KESSLER: You know what? The facts are too confused here. I'm just going to drop this subject. I don't think it's significant to the overall analysis. I have no further questions.

### 3.2 *Issue 2. Warming Transients*

The other major error in Exponent’s technical analysis pertained to the calculation of warming transients during the half-time intermission. After the footballs were returned inside at half-time to warmer temperatures (said to be about 73 deg F), they would warm relatively rapidly, with their pressures returning to their original pressure. Exponent separately showed cases in which footballs were supposedly set to 12.5 and 13 psi with the Non-Logo Gauge (Figures 24, 25 and Figure 28 left panel) and with the Logo Gauge (Figures 26, 27 and Figure 28 right panel), but actually used the Master Gauge to set these pressures to 12.5 and 13 psi, with a dramatic impact on the transients.

#### 3.2.1 *Exponent Figures 25, 27 and 28*

Exponent’s statements that it set pressures using the Logo Gauge were unambiguous. On page 52, describing the simulations depicted in Figure 27, Exponent stated that the “Logo Gauge was used to set the pressure” of Patriot and Colt balls to 12.5 psi as follows:

**Using the Logo Gauge Pre-Game:** In recognition of the remaining uncertainty as to which gauge was used to measure the footballs pre-game and in the interest of completeness, similar tests were run using the Logo Gauge. **The Logo Gauge was used to set the pressure of two balls to 12.50 psig (representative of the Patriots) and two balls to 13.00 psig (representative of the Colts).** From each set (corresponding to each team), one ball remained dry while exposed to the game temperature and the other was wet. [my bold]

The captions to Figures 26 and 27 also stated that the footballs in the simulations depicted in these figures had been “set with the Logo gauge”:

Figure 26. Transient pressure curves for the scenario in which **the footballs were set with the Logo Gauge**

Figure 27. Transient pressure curves for the scenario in which **the footballs were set with the Logo Gauge** with the average measured Game Day pressure (with error band) is overlaid.

Corresponding language was used in relation to Figures 24 and 25, which stated that the footballs depicted in these simulations had been “set with” the Non-Logo Gauge. Later in the Report, Exponent described “Game Day” simulations (p. 58), once again stating that the Logo Gauge was used to set the ball pressures pre-game.

Exponent’s Figure 28 (shown below as Figure 4) summarized both Logo and Non-Logo simulations, was discussed at length in the Appeal hearing (see pages 382-7, 403-5, 445-6.)

For the Non-Logo initialization case (Figure 28, left panel; also Figure 25), Exponent argued that there was no intersection between observed pressure (depicted as the horizontal line with the solid line at the mean of the observations and the envelope representing a confidence

interval) and the simulated transient (the lines trending upward and to the right.) The solid line of the envelope shows the transient for dry balls and the dashed line the transient for wet balls. The difference between wet and dry transients is inconsistent (for no obvious reason), with the difference ranging up to ~0.5 psi.

In the critical Logo case (Figure 28, right panel; also Figure 27), Exponent noted a very short interval at the beginning of the half-time intermission in which observed Patriot pressures overlapped with the transient. From other information, it was known that measurement of 11 Patriot footballs took at least 5 minutes, from Exponent deduced that the transients did not overlap with observations during a plausible window. Exponent accordingly declared that there was “unexplainable” deflation even in the more troublesome Logo Gauge scenario. Combining the two cases, Exponent declared that there was “unexplainable” deflation, regardless of which gauge was used by referee Anderson to measure football pressure pre-game.

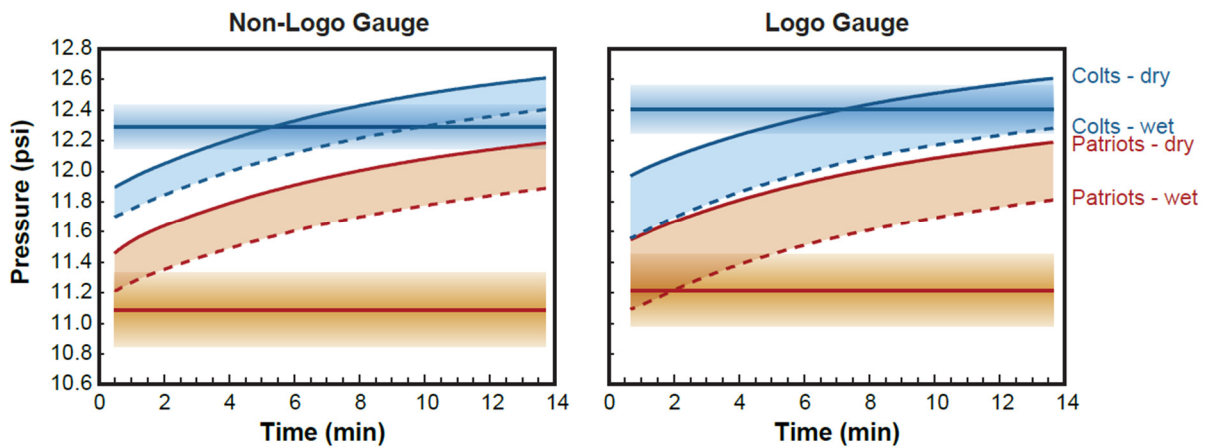


Figure 4. Exponent’s Figure 28. Original caption: A comparison of the transient pressure results from the Non-Logo (left) and Logo Gauge (right) pre-game scenarios

### 3.2.2 Re-Stated Figures 25, 27 and 28

However, these simulations were not done as described. Bizarrely, footballs in the simulations labeled “Logo Gauge” were not set to 12.5 psi with the *Logo Gauge*, but were set to 12.5 psi with the *Master Gauge* – a procedure that resulted in the footballs receiving an additional ~0.33-0.38 psi of pressure. I illustrate the difference in Figure 5, which re-states Exponent’s Figure 28 to show the transients, based on pressures set with the Logo and Non-Logo Gauge as stated in the Exponent Report (and, obviously, the relevant pressures for the simulations).

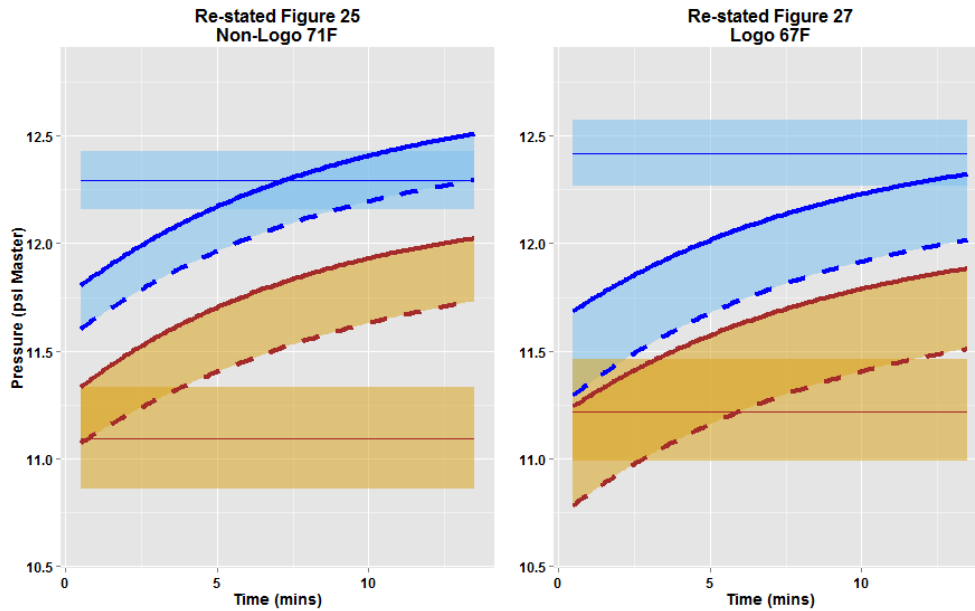


Figure 5. Re-statement of Exponent Figure 28. Left panel – re-statement of Exponent Figure 25, based on initialization using the Non-Logo Gauge (not the Master Gauge). Right panel – re-statement of Exponent Figure 27, based on initialization using the Logo Gauge (not the Master Gauge). Note that pre-game temperatures in the two Figure 28 panels are different: 71 deg F for the Non-Logo simulation and 67 deg F for the Logo simulation.

The effect on the Logo Gauge case (right panel) is dramatic: the transients move significantly downward, such that there is now a substantial window of time in which observed and transient values of Patriot balls are consistent, contradicting Exponent. In the left panel illustrating initialization with the Non-Logo Gauge, there is a more than adequate window of time in which observations and transients for Colt balls are consistent.

For reasons that are not entirely clear, Exponent’s transient diagrams for Non-Logo and Logo gauges (Figure 25 and 27) were done at different pregame temperatures (71 and 67 deg F) respectively. In Figure 6, the transient diagrams are therefore re-stated to a consistent temperature of 69 deg F, the midpoint of the estimated 67-71 deg F range for pre-game temperature. The left panel (Non-Logo) is very similar to Exponent Figure 25: the two varied effects (Non-Logo rather than Master Gauge initialization; 69 deg F pre-game rather than 71 deg F) offset one another. However, the middle panel (Logo) is even more dramatically removed from Exponent Figure 27: both transients move even lower, yielding a long window in which Patriot observations and transients are consistent.

Finally, in the right panel below, I’ve shown the case indicated by the above analysis of ball preparation as the “most likely”: Non-Logo initialization for Colt balls and Logo initialization for Patriot balls. This scenario has no discrepancies: there are lengthy intersections between observations and transients at appropriate periods for both Patriots and Colts and no indication whatever of “unexplainable” deflation.

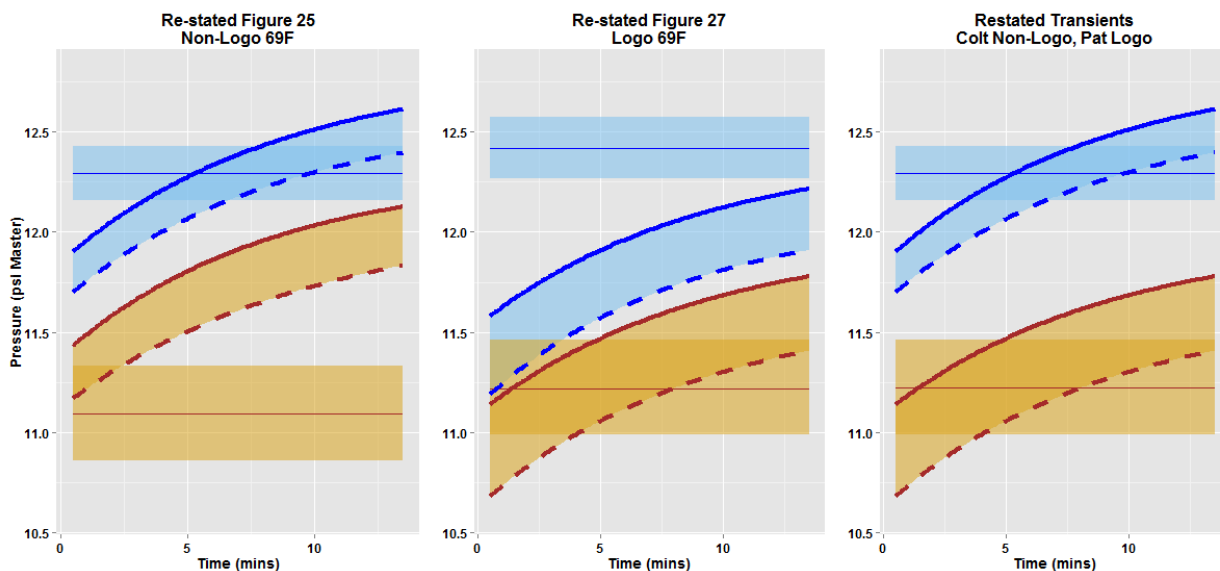


Figure 6. Left: Re-statement of Figure 25 (also left panel of Figure 28) showing results with Non-Logo initialization at 69 deg F, the midpoint of estimated pregame range, selected to be consistent in both panels (unlike Exponent). This is very similar to Exponent Figure due to offsetting adjustments. middle: Re-statement of Figure 27 (also right panel of Figure 28) showing results with Logo initialization at 69 deg F. Right: showing transients, basis 69 deg F, for Colt balls set to 13 psig using Non-Logo Gauge and for Patriot balls set to 12.5 psig using the Logo Gauge. The differences between observed values appear to come from inconsistencies in Exponent’s adjustments from Logo and Non-Logo to master gauge.

### 3.2.3 The Horns of a Dilemma

During the Appeal Hearing, Snyder had pointed out that footballs in the Logo simulations had actually been set to 12.5 psi with the Master Gauge. Caligiuri conceded that pressures had been set to 12.5 psi with the Master Gauge (*Hearing*, 374:11), but surprisingly denied that this was an error:

Q... Second, do you agree, you saw the criticism number 3 that Dr. Snyder presented, and he indicated that you should have recalibrated the starting pressures through the master gauge because you were comparing those starting pressures to the halftime pressures, which you did to the master gauge. Do you agree that you should have done that?

A. No.

Although not followed up by Brady’s counsel, Caligiuri’s denial, even if valid (which it isn’t), simply moved the error elsewhere, leaving Exponent was on the horns of a dilemma. Either the Exponent Report had incorrectly described the procedure in the simulations or the procedure in the simulations illustrated in the Exponent Report was wrong. Either way, an important correction to the Exponent Report was called for.

A precedent for correct handling of this sort of situation arose in connection with Gergis et al (2012), an academic article accepted by the *Journal of Climate*. Its authors had stated that they had screened data after de-trending, whereas, shortly after distribution of a preprint, it was shown that that they had screened data before detrending. This posed a conundrum for John Chiang, editor of the article for the *Journal of Climate* (Freedom of Information Request, 2012). The authors wanted to alter the description of methodology to match what they had actually

done. However, the editor, recognizing that peer reviewers might actually have been influenced by the stated methodology, required the authors to re-do results with the stated methodology and demonstrate that their results were not sensitive to the difference in methodology. In the end, the authors were unable to do so, and the article disappeared, somehow being annulled without ever being formally retracted.

The moral for Exponent (and Marlow) ought to have been identical. When they learned/recognized that the actual procedure for Logo simulations did not match the procedures described in the Exponent Report, they ought to have withdrawn the published version of the Exponent Report and re-issued a version in which the actual procedure for these simulations matched the originally stated procedures. If they strongly felt that the originally stated procedures were inappropriate (and, in my opinion, the originally stated procedures were fine), then they had the option of including an additional diagram.

Had this happened, Exponent would have issued a re-stated diagram along the lines of my Figures 4 and/or 5 above. Such a diagram would have demonstrated the substantial window in the Logo case during which observed pressures were consistent with modeled transients, raising a visible challenge to the key conclusion of the Wells Report that there was no such window *regardless* of gauge used to measure Patriot ball pressure.

## 4 Discussion

In this section, I will discuss several ethical and policy issues arising from the above analysis of the controversy:

- Were the Patriots culpable anyway?
- Was anyone negligent in the NFL's technical analysis and, if so, who?
- How to fix the situation
- The rhetoric of "science says"
- Lack of common sense

### 4.1 *Were Patriots Culpable Anyway?*

The above analysis shows that it is "very likely" that the Patriot footballs were under-inflated footballs when tendered to the referee. Should they be punished anyway? Should Brady be suspended anyway?

First and most obviously, this becomes an entirely different case than the accusation of McNally "tampering" with footballs in the washroom after measurement by the referee. McNally had no role in Jastremski's preparation of footballs or Jastremski's decision to set the pressures to 12.6 psi after rubbing, rather than before. Accordingly, McNally's emails become moot. On their own, the emails had other interpretations which had been rejected by Wells based on a deflation narrative involving McNally, but Wells' adverse interpretations are simply moot.

Second, there is no evidence that the Patriots, Brady or even Jastremski himself realized that setting pressure after gloving would result in under-inflation of ~0.35 psi at room temperature. The failure of the Patriots themselves to adduce this argument is eloquent testimony. In addition to the Patriots, one can add Exponent and its scientists, Daniel Marlow,

the lawyers of Paul, Weiss and the NFL front office to those who failed to appreciate that Jastremski's technique would result in under-inflation.

Third, although the Wells Report misleadingly stated that "the Patriots have developed a process for the preparation of game balls in accordance with the preferences of Tom Brady, who has been the team's starting quarterback for over thirteen years" (*WR*, 17), the ball preparation technique used for the AFC Championship had not been used in the previous few seasons during Jastremski's tenure as equipment manager. Both the Wells Report (*WR*, 49ff) and Brady's consistent evidence at the Hearing (*Hearing*: 55, 68-70, 76-77) demonstrate that Brady was worried about the texture of wet balls and asked Jastremski to prepare the footballs with manual gloving, rather than Lexol, a technique that had been used by Jastremski's predecessor but not by Jastremski. To the extent that Jastremski's procedure resulted in slightly under-inflated footballs being tendered to referee Anderson, it was not due to any long-standing Patriot subterfuge.

Indeed, had Jastremski previously used this particular technique, he almost certainly would have learned about the effect of setting pressure after gloving. Because the procedure necessarily results in footballs being under-inflated at room temperature, the slight under-inflation would have been picked up in any earlier circumstance by a referee with an accurate gauge.

Fourth, there is no evidence that Jastremski attempted to deceive the referees through his technique. While he presented balls that were slightly under-inflated at room temperature, he could hardly have expected that the NFL official responsible for a championship game would use a gauge that was biased by a similar amount.

Fifth, to the extent that additional deflation of ~0.35 psi was a competitive concern, the Patriots could hardly have expected that referee Anderson would have had a gauge that read ~0.38 psi too high. In Belichick's press conference of January 24, 2015, he seemed to think that the referees set the ball pressure anyway.

## 4.2 *How to Fix*

Without the scientific errors in the Exponent Report, it is doubtful there would be any Deflategate controversy. Unfortunately, the original errors were compounded by the failure of Brady's lawyers and technical consultants to either identify (simulation of ball preparation) or effectively argue against ("Logo Gauge" simulations) the errors. Critics of Commissioner Goodell generally overlook his awkward position: he had received an opinion adverse to the Patriots from reputable and competent scientists which was weakly opposed.

Appeal courts are singularly ill-suited to redressing this sort of error. They are reluctant to substitute their judgement on issues of fact, and, as a result, appeal processes are almost always about issues of procedure and fairness, rather than facts, with the present litigation being an example. Worsening Brady's position in the present appeal is the fact that recent amicus briefs have presented similar arguments from those presented by Snyder at the Appeal Hearing, without showing errors on Exponent's part.

There is another way to resolve the controversy.

The scientific community takes considerable pride in the concept of science being "self-correcting". Under this philosophy, it is the scientific community, not appeal courts, which has

responsibility for identifying scientific errors and correcting the scientific record. When a scientist has inadvertently made an error, the most honorable and effective method of correcting the scientific record is to issue a corrected report, and, if such is not possible, retraction. If either Exponent or Marlow conceded the above errors, it is hard to envisage the Deflategate case continuing much further. Accordingly, even at this late stage, Exponent and/or Marlow should man up, acknowledge the errors and either re-issue corrected reports or retract.

While academics are often reluctant to admit error, in the present case, both Exponent and Marlow have duties not only to their immediate client (the NFL), but to Brady and the Patriots, as third parties, a duty implicitly recognized by Caligiuri (Hearing, 348:11). It is surely arguable that the errors identified in this article are ones that ought not to have been made by professionals with the high level of skill attributed to Exponent and Marlow. Be that as it may, they surely have a continuing duty to Brady and the Patriots to acknowledge errors as soon as they become aware of them.

In respect to the erroneous simulation of ball preparation, there is a curious uncertainty about whether Exponent was aware of the detail that Jastremski set pressure after gloving and overlooked it in their analysis or whether Paul, Weiss forgot to inform Exponent of this detail. Exponent carefully said that its analysis was based on information from Paul, Weiss and, in most cases, descriptions in the Wells Report are consistent with descriptions in the Exponent Report. The inconsistency in this instance is rather striking. If Exponent was unaware of this detail, they should disclose that as part of their correction.

Out of all the procedural issues raised in the appeal, the withholding of instructions from Paul, Weiss to Exponent seems particular invidious to Brady's ability to respond to Exponent's scientific conclusions. This issue was forcefully raised in the Patriots amicus brief as follows:

In short, the Commissioner relied on the Wells Report. The Wells Report relied on Exponent's "conclusion" that science did not explain the PSI of the Patriots footballs. Exponent based that conclusion on assumptions from Paul Weiss. Those assumptions could only be tested by having access to the interview notes sought in discovery. The Commissioner refused to allow that discovery.

Another important instruction that might have been mentioned in this context was the instruction in which Exponent was "told" not to consider the possibility that Anderson switched gauges. Had Brady's technical consultants been able to consider such instructions, their analysis of the Exponent Report might well have been more effective.

### **4.3 The Rhetoric of "Science Says"**

Re-reading Wells' eloquent testimony at the Appeal Hearing, it is striking how he used the rhetoric of "science says" to support arguments that were weakly connected to "science" and which were ultimately incorrect. For example, Wells stated that "scientific tests" supported the finding that the Non-Logo Gauge was used (*Hearing*, 301:14):

The scientists, the Exponent people say they believe based on their **scientific tests** that the non-logo gauge was used.

Wells also claimed that "science" entitled him, as finder of fact, to disregard Anderson's recollection of having used the Logo Gauge, grandiosely equating Exponent's analysis of gauges to a "tape" of events (*Hearing*, 303:17):

Q. Okay. So in your role as the judge, okay, you concluded that you were going to reject as a finder of fact Mr. Anderson's best recollection that he used the logo gauge, correct?

A. Not only did I reject it, I first said this is what he says and this is why I am rejecting it. And I set it out so everybody can see it. Look, this is no different than a case where somebody has a recollection of X happening and then you play a tape and the tape says Y happened. Now, the person could keep saying, well, darn it, I remember it was X. But the people are going to go with the tape. I went with the science and the logic that I had three data points. And that's what I based my decision on.

In reality, Exponent had not done any “scientific tests” to show that the Non-Logo Gauge had been used. Rather, they had deduced that Anderson had used the Non-Logo Gauge from circumstantial evidence (the similarity of gauge readings). As it turned out, there was another available interpretation of the circumstantial evidence that they overlooked and their interpretation of the circumstantial evidence was incorrect.

Advocate and even policy-makers often like to say that their conclusions are supported by “science”, but Wells’ enthusiastic use of the terms “science” and “scientific tests” as rhetoric to validate incorrect analysis should serve as a caveat.

## 5 Conclusion

There have been many insightful criticisms of the NFL’s Deflategate case, nearly all of which have focused on the unfairness of the process and uncertainty of the science. However, none of the major critiques or briefs successfully confront the scientific basis (endorsed by Wells and the NFL) used by Exponent in their finding that referee Anderson used the Non-Logo Gauge to measure Patriot footballs pre-game, a finding that is essential to the NFL’s case. Their basic strategy is to argue that there is uncertainty about which gauge was used to measure Patriot ball and that the uncertainty is too great to support the NFL’s punishment.

There is an alternate approach. In their simulations of Patriot ball preparation, Exponent made a basic error which has missed by all commentators to date (including my own prior commentary) and which, when corrected, resolves the entire affair, fully explaining the “unexplained” deflation, while demonstrating that there was no “tampering” with the footballs after measurement by the referees.

The newly identified error pertains to Exponent’s simulations of Patriot ball preparation, an issue identified by Bill Belichick in his first press conference, but dismissed by Exponent and the Wells Report. In Exponent’s simulations, they set football pressure to 12.5 psi before gloving, whereas the Wells Report reported that Patriot equipment manager Jastremski set pressure to 12.6 psi **after** gloving. In Exponent’s simulations, gloving during ball preparation increased temperatures and pressures by about 0.7 psi, but the effect wore off in 15-20 minutes. However, Jastremski’s actual technique necessarily resulted in Patriot balls being slightly under-inflated by about 0.35 psi (approximately 12.1-12.2 psi) when they returned to room temperature. This amount fully accounts for the “unexplained” additional deflation of Patriot balls.

This under-inflation, while slight, was still enough that it should have been observed by referee Anderson in his pre-game measurement. However, Referee Anderson had two gauges, one of which (the Logo Gauge) read 0.38 psi too high. By coincidence, the amount of under-inflation under Jastremski’s protocol closely matched the bias in Anderson’s Logo Gauge. Ten of 12 Patriot balls were measured by Anderson between 12.5 and 12.6 psi, with two under-



inflated. This is only possible if Anderson used the Logo Gauge to measure Patriot balls, resolving a battleground issue contrary to the findings of Exponent, Wells and Goodell. This is consistent with Anderson's recollection of having used the Logo Gauge, a recollection rejected by Exponent and Wells.

On the other hand, Exponent's reasoning in respect to Colt balls, the pressures of which were set at room temperature without rubbing, remains valid: Anderson used the Non-Logo Gauge to measure Colt balls. The corollary is that Anderson (inattentively) switched gauges between measuring Patriot and Colt balls. An excellent precedent for this possibility is the identical inattentive switch of gauges by NFL officials at half-time, even under heightened scrutiny.

One of the conclusions of the present analysis is that there is a coincidence between the amount of under-inflation below 12.5 psi arising from Jastremski's ball preparation and the bias of Anderson's Logo Gauge, but it is a coincidence that can be documented. At the Appeal Hearing, Wells spoke eloquently against coincidence ("lightning strike"), but his own theory ultimately rests on an implausible coincidence. Wells' theory requires that, out of all possible deflations available, the Patriots decided to deflate their footballs by the amount of bias of referee Anderson's Logo Gauge. Wells should have been worried about his own lightning strike.

The Deflategate controversy originated in scientific and technical errors. Appeal courts are poorly suited to resolve such errors. There is another way to resolve the controversy. The scientific community takes considerable pride in the concept of science being "self-correcting". Under this philosophy, it is the scientific community, not appeal courts, which has responsibility for identifying scientific errors and correcting the scientific record. When a scientist has inadvertently made an error, the most honorable and effective method of correcting the scientific record is issue a corrected report, and, if such is not possible, retraction. If either Exponent or Marlow conceded the above errors, it is hard to envisage the Deflategate case continuing much further. Accordingly, even at this late stage, Exponent and/or Marlow should man up, acknowledge the errors and either re-issue corrected reports or retract.

## 6 References

Blecker, Robert (2015, Dec 14). Brief for Professor Robert Blecker as Amicus Curiae in support of Appellees. Downloaded from <http://thewhitebronco.com/wp-content/uploads/2015/12/Deflategate-Amicus.pdf>

Exponent (2015, May 6). The Effect of Various Environmental and Physical Factors on the Measured Internal Pressure of NFL Footballs. In Wells et al, Investigative Report Concerning Footballs Used during the AFC Championship Game on January 18, 2015.

Freedom of Information Request (2012). Part 2(a), Correspondence between paper authors and Journal of Climate. Retrieved from <http://www.climateaudit.info/correspondence/foi/gergis/Part%202a%20Journal%20Correspondence.pdf>

Glanz, James. (2015, January 29). Deflation Experiments Show Patriots May Have a Point After All. *The New York Times*. Retrieved from

<http://www.nytimes.com/2015/01/30/sports/football/deflation-experiments-show-patriots-may-have-science-on-their-side-after-all.html>

Hassett, Kevin A., Joseph W. Sullivan and Stan A. Veuger, (2015, June 12). On the Wells Report. *AEI Economic Perspectives*. Downloaded from <https://www.aei.org/wp-content/uploads/2015/06/On-the-Wells-report.pdf>

MacKinnon, Roderick, (2015, May 14). Professor MacKinnon's Scientific Conclusion. Retrieved from <https://wellsreportcontext.com/mackinnons-scientific-conclusion/>

National Football League, (2015, June 23). In the Matter of Thomas Brady: Appeal Hearing Transcript. Downloaded from <http://static.nfl.com/static/content/public/photo/2015/08/04/0ap3000000506502.pdf>

National Football League (2015, July 23). Final Decision on Article 46 Appeal of Tom Brady. Downloaded from <https://nflabor.files.wordpress.com/2015/07/07282015-final-decision-tom-brady-appeal.pdf>

National Football League Management Council v National Football League Players Association and Tom Brady (2015, Sept 3). Decision and Order. 15 Civ 5916 (RMB)(JCF), 15 Civ 5982 (RMB)(JCF). Downloaded from <https://assets.documentcloud.org/documents/2330130/judge-dismisses-tom-bradys-four-game-suspension.pdf>

NFL Staff, (2015, Sept 3). Tom Brady Suspension Timeline. Downloaded from <http://www.nfl.com/news/story/0ap3000000492189/article/tom-brady-suspension-timeline>

New England Patriots (2015, May 14). The Wells Report in Context. Retrieved from <https://wellsreportcontext.com/>

New England Patriots (2016, May 25). Brief for New England Patriots, L.L.C. as Amicus Curiae supporting the petition of Appellees National Football League Players Association and Tom Brady for Rehearing or Rehearing en Banc. Retrieved from [http://cdn-s3.si.com/s3fs-public/download/patriots\\_accompanying\\_brief\\_for\\_the\\_amicus\\_filing.pdf](http://cdn-s3.si.com/s3fs-public/download/patriots_accompanying_brief_for_the_amicus_filing.pdf)

Professors of Physics and Engineering, (2016, May 24). Brief of Professors of Physics and Engineering as Amici Curiae in support of Appellees' Petition for Panel Rehearing or Rehearing en banc. Retrieved from <https://cbsboston.files.wordpress.com/2016/05/physics-professors-deflategate-filing.pdf>

Troy Vincent's Letter to Tom Brady (2015, May 12). Retrieved from [http://espn.go.com/nfl/story/\\_/id/12873455/troy-vincent-letter-tom-brady](http://espn.go.com/nfl/story/_/id/12873455/troy-vincent-letter-tom-brady)

Wells, Theodore V. Wells, Jr., Brad S. Karp and Lorin L. Reisner. Paul, Weiss, Rifkind, Wharton and Garrison LLP. (2015, May 6). Investigative Report Concerning Footballs Used During the AFC Championship Game on January 18, 2015. Downloaded from <https://nflabor.files.wordpress.com/2015/05/investigative-and-expert-reports-re-footballs-used-during-afc-championsh.pdf>