CHAPTER 12

Critical Thinking as Scientific Reasoning
Examining the Power of Sports Momentum

John Ruscio and Kevin Brady
The College of New Jersey

Introduction

At the start of the 2016 season in the National Football League (NFL), the Philadelphia Eagles won their first three games. Under the leadership of new head coach Doug Pederson and with the promising talent of rookie quarterback Carson Wentz, they seemed unstoppable. There was endless discussion in the sports media about how far the Eagles' momentum might carry them. What nobody saw coming is that they would lose 9 of their remaining 13 games, ending the season with a 7–9 record and missing the playoffs. Just as the Eagles' decline was beginning, enthusiasm shifted to the Minnesota Vikings. They opened their season even stronger, winning their first five games. This, it seemed, was the team that had all the momentum and would take the league by storm. The Vikings proceeded to lose eight of their remaining eleven games, ending the season with an 8–8 record and missing the playoffs. Something like this happens every year in the NFL, and it is by no means unique to that league, that sport, or any sport. A string of victories creates the perception of momentum and an expectation of continued success. When win streaks come crashing down to Earth, does this diminish belief in the power of sports momentum? Not a chance!

Back up a couple of years, when the New England Patriots opened the 2014 season with a 2–2 record. With the legendary combination of coach Bill Belichick and quarterback Tom Brady, the Patriots had been on an unprecedented run of success for more than a decade and were perennial contenders for a Super Bowl appearance. After an embarrassing loss to the Kansas City Chiefs in week 4, the sports world could talk of little else but "the end of the dynasty". Their momentum had finally run its course, and there were calls for replacing Belichick and benching Brady. At the same time, the Seattle Seahawks, coming off a resounding Super Bowl victory over the high-octane Denver Broncos, started their season with a 3–3 record. They, too, appeared to have lost all their momentum. Despite their falls from exalted status, both teams would end the season with records of 12–4, earning the top playoff seeds in their respective conferences. They battled their way through tough playoff games to earn spots in the Super Bowl, which was another hard-fought contest that came down to the final seconds of the game, as the Seahawks opted not to give their star running back a chance to win the game from one yard out and Malcom Butler sealed the victory for the Patriots with a goal-line interception. Do outstanding finishes to seasons that begin with bitter disappointment diminish belief in the power of sports momentum? Again, not a chance.

How is it that a belief can survive in the face of apparently unsupportive evidence? Where do beliefs of questionable merit come from in the first place, and how do they gain traction? Rather than gullibly believing every new idea that comes along or close-mindedly rejecting them all, what can we do to maintain an open-minded skepticism, to think critically about claims? The literature on judgment and decision making (Kahneman, 2011) documents the many shortcomings in our thinking that make us susceptible to erroneous beliefs, and the tools of scientific reasoning provide us with means of protection (Sagan, 1995). The case of sports momentum will be used to illustrate our cognitive vulnerability as well as the role of critical thinking in arriving at more accurate conclusions.

What is Sports Momentum?

Michael Kent (2006) defines sports momentum as:

> The positive or negative change in cognition, affect, physiology, and behavior caused by an event or series of events that affects either the perceptions of the competitors or, perhaps, the quality of performance and the outcome of the competition. Positive momentum is associated with periods of competition, such as a winning streak, in which everything seems to "go right" for the competitors. In contrast, negative momentum is associated with periods, such as a losing streak, when everything seems to "go wrong" (Kent 2006, p. 444).

There are two important parts to this definition. First, sports momentum is said to consist of changes in thought, feeling, or behavior brought about by past performance. For example, an athlete might become more confident in his or her ability to perform well as a result of recent success, or less confident as a result of recent failure. Even if not at a professional level, and even if not in the realm of sports, we have all experienced thoughts and feelings associated with successful (as well as unsuccessful) performance at a task. There is no denying the existence of such psychological experiences. This component of sports momentum, the psychological changes caused by past performance, is not in doubt.

The second part of the definition, on the other hand, is where greater skepticism may be warranted. It is said that the psychological changes experienced in the wake of success or failure influence *either* how athletes are perceived *or* how they perform. This is a crucial distinction. It acknowledges that belief in the power of sports momentum to affect performance may be a cognitive illusion, a perception that is not matched by reality. Is sports momentum merely descriptive, identifying a run of success (or failure), or does it hold predictive value, signaling an increased likelihood of continued success (or failure)?

Most people who talk about sports momentum appear to ignore the possibility that it is a cognitive illusion and believe instead that it holds predictive value – that there is a causal mechanism by which past performance affects future performance. For example, an NFL team that has won several games in a row is thought to bring something to their next game that makes them a tougher opponent than a team with a comparable win–loss record but not a run of recent success. Why does it seem that we are primed to believe in the power of sports momentum? Why does the possibility that it might be a cognitive illusion seldom receive serious consideration? Why does belief in the power of sports momentum persist despite abundant evidence that challenges it? The answers to these questions, along with others regarding a wide range of questionable beliefs, can be found through an exploration of the cognitive heuristics, or mental shortcuts, that predominate when we are engaged in uncritical thinking, which is our default state.

Uncritical Thinking

Kahneman (2011) describes a metaphor for human thinking that encompasses two different systems, which for convenience he labels as System 1 and System 2. Whereas System 2 is effortful, slow, and requires deliberate attention and control, System 1 is comparatively effortless, fast, and automatic, operating mostly outside conscious awareness. Just as it would be impossibly taxing on the brain's limited resources to pay conscious attention to the processing of all sensory input that it receives, it would be impossibly taxing to engage System 2 to reach every judgment or decision in our lives, to evaluate every claim to knowledge that we encounter. Out of practical necessity, most of our thinking takes place within the realm of System 1.

The mental shortcuts that operate in System 1, which have come to be known as cognitive heuristics (Tversky & Kahneman, 1974), generally serve us well, especially in terms of their efficiency. They allow us to reach judgments or decisions quickly and easily, and particularly when we consider information well within the confines of our everyday experience. System 1 usually steers us in the right direction. There is, however, a "satisficing" quality to much of what takes place in System 1 (Simon, 1956), an implicit criterion of "good enough" output that sacrifices some accuracy in order to maintain cognitive ease. System 2 contains tools capable of greater accuracy, but because it is not feasible to employ these for most of our thinking, we rely heavily on System 1. The tools of critical thinking reside almost exclusively in System 2. Most of these need to be learned through study, honed through practice, and applied with considerable effort when that seems worthwhile. Because the shortcuts in System 1 operate in an effortless, fast, and automatic manner, this will be our default mode of uncritical thinking.

If we must use mental shortcuts operating outside conscious awareness to avoid mental exhaustion, how do we come to understand these heuristics? Another analogy with perception holds the answer. Just as perceptual illusions can teach us a lot about the ways that our perceptual systems work, cognitive illusions likewise teach us a lot about the heuristics in our System 1 repertoire. In a seminal paper synthesizing much of their earliest work together, Tversky and Kahneman (1974) described three heuristics they discovered by observing the systematic and predictable mistakes they and their research subjects made in a variety of tasks.

The Representativeness Heuristic

One mental shortcut of System 1 is the representativeness heuristic with which we estimate probability or likelihood according to perceived similarity (Kahneman & Tversky, 1972). For example, try rank-ordering the likelihood of observing the following three ordered sequences of girl (G) and boy (B) births: (a) GGGGBBBB, (b) GBGBGBGB, (c) GBBGGGBB. Sequence (a) has an obvious pattern that makes it seem unlikely to occur by chance, and to a lesser extent so does sequence (b). Sequence (c), by comparison, is more similar to – or more representative of – what we expect by chance. Comparing the similarity of an observation to an expectation is not fundamentally unreasonable, and it certainly is an efficient way to reach a judgment. This is especially true when a more objective approach is either unavailable (e.g., we don't know how to calculate probabilities) or not feasible (e.g., we don't care to spend the time to do so).

When engaged in System 1 thinking, we rely on the representativeness heuristic to avoid the harder work of thinking in terms of objective probabilities, substituting instead an easier task that can be handled quickly. Rather than calculating the probability of each G-B sequence, we form an impression of which sequences are more similar to what we expect randomness to look like. Patterns of any kind do not appear random, and therefore they seem unlikely to occur by chance. This cries out for an explanation. Of course, all randomly generated G-B sequences of the same length are equally likely to occur (given the assumption that 50 percent of all births are girls and 50 percent are boys), even those as seemingly non-random as GGGGGGGG and BBBBBBBB.

Belief in the power of sports momentum probably benefits from the representativeness heuristic. A winning (or losing) streak is inconsistent with what we expect to observe by chance. Of course, it's well known to psychological scientists that people are biased in their attempts to produce or recognize truly random sequences (Nickerson, 2002). We intuitively expect to see fewer clusters, and more alternations, than actually occur by chance. Philadelphia has won three games in a row; the Vikings have won five in a row: clearly these teams have the momentum on their side. That's as far as the representativeness heuristic is likely to take us.

The Availability Heuristic

A second shortcut of System 1 is the availability heuristic with which we estimate frequency or probability according to the ease with which instances can be recalled (Tversky & Kahneman, 1973). For example, do you think that more people die from homicide or suicide in the United States each year? Because few of us are familiar with objective data on the frequency of various causes of death, and unlikely to spend time looking up the facts unless there is a compelling reason to do so, we rely on the availability heuristic. This substitutes a question that can be answered fairly easily: Which comes to mind more readily, instances of homicide or suicide? Whichever we can recall more easily will be judged more common.

The availability heuristic tends to work well, as things that are more frequent in our environment usually do leave behind more instances in our memory and are therefore easier to recall. Biases creep in when some events are over- or underrepresented in our memories – due to their vividness, salience, danger, or media coverage, for example – and thereby become artificially easier to recall. Homicides generally attract much more media coverage than suicides, bringing them to the attention of a larger audience. For a variety of reasons, suicides tend to remain comparatively private. Whereas many people believe homicide to be a more common cause of death than suicide, comprehensive data show the reverse. Annual reports from the Centers for Disease Control show that in a typical year in the United States, there are more than twice as many deaths by suicide than by homicide.

Belief in the power of sports momentum probably benefits from the availability heuristic. A winning (or losing) streak is much more salient and vivid than a mixed series of wins and losses. These streaks receive considerable attention in sports media. This enhances our likelihood of being exposed to such streaks, often repeatedly, and thereby of encoding them for later retrieval from memory. When we do find ourselves estimating their frequency, the availability heuristic is likely to yield an overestimate due to our artificial overexposure to streaks. The perceived frequency of streaks supports a belief in the power of sports momentum.

The Anchoring and Adjustment Heuristic

A third shortcut of System 1 is the anchoring and adjustment heuristic with which we estimate an unknown quantity by starting with an initial value and moving in the correct direction, but not far enough (Tversky & Kahneman, 1974). For example, take five seconds to estimate the following product:

$$1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$$

It is highly unlikely that you could do this arithmetic so quickly. Multiplication is a perfect example of a deliberate, effortful, and slow process that takes place in System 2 thinking. We must work hard to learn the skill, and we must commit effort to use it. Even with more time, most people would find this to be a very challenging problem to solve in their head.

By using System 1 thinking, the anchoring heuristic enables us to simplify the problem to meet the constraints of time and cognitive abilities. You would probably begin with some multiplication. Suppose that five seconds is enough for you to recognize that 1×2 is 2, 2×3 is 6, and 6×4 is 24. As time expires, what you have is a starting point, or anchor. You know that the correct answer to the problem is higher still, because there are more numbers left to multiply, so you adjust upwards. In Tversky and Kahneman's (1974) original study, university students' median estimate was 512 – but the correct answer is 40,320. A partial calculation provides an anchor, but the adjustment does not go nearly far enough in the right direction.

Further evidence that a partial calculation serves as an anchor comes from a variant of the problem. Suppose you had been given five seconds to estimate this product, instead:

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

If you begin with some multiplication, five seconds might be enough for you to figure out that 8 × 7 is 56, and 56 × 6 is, well, something greater than 300. Look what has happened to the anchor. Rather than 24, now it is already over 300. When given this version of the problem, students in Tversky and Kahneman's (1974) study gave a median answer of 2,250. It is still much too small, but it is more than four times larger than before. The arbitrary choice of which order to write the numbers in the problem changed the anchor, which in turn affected the final estimates. Research has demonstrated that anchors can be obtained in many ways, including transparently arbitrary methods such as random spins of a wheel or reference to digits in a personal ID number or less obviously biased methods such as partial calculations or retrieval of memories (Furnham & Boo, 2011). Whatever the source of an anchor, the adjustment tends to be insufficient in magnitude.

Belief in the power of sports momentum is likely to benefit from the anchoring and adjustment heuristic. The anchor in this instance is recent experience, either personal or vicarious, involving a winning (or losing) streak. An estimate of future performance is not adjusted sufficiently far from this anchor of extremely good (or poor) performance. Even though we recognize it is unlikely the Eagles or Vikings will remain undefeated through a full sixteen-game season, we fail to adjust far enough away from their early run of success and end up surprised by the more ordinary record that follows. This also constitutes a failure to understand regression to the mean, another concept that is entirely foreign to the cognitive heuristics of System 1.

Other Characteristics of System 1

The representativeness, availability, and anchoring and adjustment heuristics are not the only ways that System 1 helps us cope with complexity and time constraints. Kahneman (2011) describes many additional characteristics of System 1, such as distinguishing the surprising from the normal, inferring and inventing causes and intentions, neglecting ambiguity and suppressing doubt, and being biased to believe and confirm. Once again, the characteristics of System 1 are apt descriptions of uncritical thinking because they promote a relatively blind acceptance of beliefs based on minimal evidence. At the same time, however, they are invaluable tools.

The ability to identify patterns and assign meaning to them almost certainly holds great survival value, as this can alert us to potential threats or opportunities and impel us to act. That some of these patterns will be cognitive illusions that prompt us to take unnecessary action demonstrates only that we may be biased to err on the side of caution. In dealing with some false-positive beliefs that turn out not to be threats or opportunities, little may be lost in the process. On the other hand, the absence of highly sensitive pattern-recognition and meaning-making abilities could be more costly, causing us to miss out on opportunities or to take defensive action in the face of real threats. It is far too cumbersome to engage in critical thinking continuously; this would literally deplete scarce mental resources (Kahneman, 2011). A default mode that excels at identifying patterns and ascribing significance to them, and that does this with great ease, is an exceptional achievement in cognitive architecture.

It turns out that a great many characteristics of System 1 help to understand how dubious ideas, such as belief in the power of sports momentum, can emerge and spread. A winning or losing streak is surprising, relative to more normal fluctuations in performance. It stands out as a pattern that needs to be explained, and the notion of sports momentum fills this void perfectly. By biasing our thinking toward belief (e.g., that past performance exerts an influence on future performance), seeking confirmation (e.g., noticing other streaks, remembering instances where streaks persisted) while suppressing ambiguity and doubt (e.g., forgetting instances where streaks ended, not worrying too much about how it is that the past performance plays a causal role in future performance, nor about the possibility that streaks themselves might occur by chance), confidence in the power of sports momentum can increase. Between the three heuristics and many additional characteristics of System 1, our uncritical thinking primes us to perceive clustered series of outcomes rather than random sequences of wins and losses, to believe that momentum is their causal impetus, and to become confident in this belief rather than submitting it to careful scrutiny or rigorous tests.

Critical Thinking

With so many forces aligning to create and bolster support for a dubious belief such as the power of sports momentum, it's clearly also of value to be capable of critical thinking. If System 1 operates primarily as uncritical thinking, we must turn to System 2 as a source of critical thinking. However, simply exercising voluntary control and devoting attention and effort to our thought process does not necessarily enable us to reach more accurate conclusions via System 2 than we would with an easy reliance on System 1. If you are unable to multiply large numbers in your head, for example, you can devote all the time you like to System 2 thinking but you will still fail to calculate correctly that $1 \times 2 \times 3 \times \ldots \times 8 = 40{,}320$.

To assist in critical thinking, System 2 must be equipped with tools that effectively combat our mental shortcomings, such as our cognitive limitations and biases. This is just what scientific methods have been designed to do, and to the extent that we can successfully learn to engage in more scientific reasoning, there is hope that we can overcome the tendencies to overidentify patterns, infuse them with too much meaning, and hold them with undo confidence. These tools do not come naturally to most people. Anyone who has studied, let alone taught, even the most fundamental principles of research design, measurement, or statistics knows how challenging it can be to master these skills. These are among the tools that must be added to the repertoire of System 2 so that they can be used to gather pertinent information and examine it systematically in order to test ideas rigorously. Thus, we define critical thinking as scientific reasoning, as equipping System 2 to use an array of human inventions specifically designed to overcome the inherent limitations and biases that lead to efficient but error-prone System 1 thinking.

Is It Worth the Effort?

Thinking outside the lab as you have learned to think in the lab can improve your judgments and decisions, and can help to evaluate the merit of claims to knowledge. Just as using System 2 for any other purpose is highly effortful, the same will be true of critical thinking. Is it worth the commitment of resources? Will it be worth the trouble to think hard about the logic and evidence bearing on a questionable belief? What if this requires gathering more comprehensive information and examining it for yourself? If critical thinking requires that new tools be added to the repertoire available to System 2, will it be worth the time and effort to learn them, to practice them, and to apply them carefully? Not everyone will be willing to do so.

In addition to cognitive skills, there are also tools that can extend the observational powers of our senses (e.g., microscopes, telescopes, infrared or ultraviolet cameras, time lapse or slow-motion video playback), the retrieval of information beyond what is accessible in our memory (e.g., archival records), or the computational power of our brains (e.g., calculators, spreadsheets for data management, software for statistical analysis). The proper use of these tools must be learned. The investment in learning to work with such hardware and software can be enormously helpful to gather and examine a wider range of relevant information, and taking advantage of the speed and power of computers allows us to perform many tasks more efficiently and accurately. These tools are not substitutes for critical thinking, they are essential components.

Whereas uncritical thinking relies on the effortless, fast, and automatic processes of System 1, critical thinking requires the will to commit time and effort to the more demanding processes of System 2. Moreover, it also requires that we have equipped System 2 with the necessary tools. Some of these tools, in turn, will require external assistance. An informative statistical analysis, for example, is unlikely to be possible without at least using a calculator plus a reference guide to supply the appropriate formulas and tables, if not computer software to enable more complex analyses of larger data sets. Often, then, there will be layers of special-purpose tools that demand effort-intensive steps to learn and to apply on a case-by-case basis.

To justify all of this effort, there must be something significant at stake. Those who aspire to a career in a scientific discipline devote enormous amounts of time and energy to learning the tools of the trade. The stakes are professional: obtaining a job, securing funding, publishing research, earning promotions, attaining honors and awards, and so forth. Nonscientists, or scientists operating outside their areas of expertise, may have considerably less incentive to acquire the skills and master the tools required to think critically about a particular claim to knowledge. When there is relatively little at stake, when there are few or no personal consequences attached to the accepting a belief, rejecting it, or remaining undecided, it should not be surprising that questionable beliefs will face only the uncritical thinking of System 1.

It is possible that belief in the power of sports momentum persists in part for this very reason – that neither holding nor rejecting this belief is very costly. Other questionable beliefs in the realm of sports have been changed in response to critical thinking when the stakes became high. For example, as Lewis (2003) described so vividly in *Moneyball*, there were significant flaws in the conventional methods used to assign value to baseball players. Scouts and coaches looked for certain physical attributes and prized certain statistical accomplishments, but it turns out that more objective analyses using improved data sources provided better measures of value. This enabled savvy teams to identify undervalued players. Lewis chronicles the remarkable success of the Oakland Athletics, who leveraged one of the lowest payrolls in all of Major League Baseball (MLB) to field teams that competed successfully with those of the highest payrolls. The driving force behind this revolution in baseball thinking was Billy Beane. When he declined an offer to become the General Manager (GM) of the Boston Red Sox in 2002, they turned instead to Theo Epstein, who became the youngest GM in MLB history. Epstein ran with Beane's ideas and achieved even greater success. In 2004, the Red Sox won their first World Series in 86 years, and they have won it three more times since then. In 2011, Epstein became President of the Chicago Cubs, and in 2016 they won their first World Series in 108 years.

Throughout the world of professional sports, it is becoming clear that bringing in talented data analysts can help to build a stronger roster as well as make better-informed coaching decisions. This revolution in decision making has happened precisely because there is so much at stake in professional sports. The same will hold true for certain college sports, too, where there is a lot of money on the line (e.g., athletes' prospects for professional careers, college coaching salaries, college revenue streams such as television revenues or fundraising). It won't be surprising to learn that data-driven strategies assist in the recruitment of talented athletes or in making game-time decisions. It would be surprising to see this approach taken in a local youth sports league, as the potential rewards simply do not justify the effort.

Something similar – the lack of significant consequences – may explain why beliefs about sports momentum are accepted with little or no scrutiny, why they're rarely submitted to rigorous tests, and why they are so resilient to challenge. The psychological experience of changes in feelings, thoughts, and behaviors following success or failure is real, and most people can relate personally to this aspect of momentum even if not as professional athletes. The question of whether these changes affect subsequent performance may not actually matter very much. Winning or losing streaks seem special, and all the forces of uncritical thinking push in the same direction, to accept the power of sports momentum.

Even if all the talk about sports momentum is grounded in nothing more than a cognitive illusion, perhaps the only people eager to know the truth would be those with something significant at stake. For example, anyone wagering on sports might take an interest in this subject. Betting lines are responsive to the beliefs of those placing bets, whether they are well-informed or otherwise. If most gamblers believe in the power of sports momentum and it turns out this is a cognitive illusion, this could create a profitable betting opportunity. The general strategy would be to bet against streaks because the betting lines may be skewed by a misconception. For example, after winning their first three games in 2016, the Eagles were favored by 3 points in their fourth game; they lost by 1. Likewise, after winning their first five games, the Vikings were favored by 3 points in their sixth game; they lost by 11. Betting against both teams at these times would have paid off. Of course, these examples were not selected at random and may not be representative. Bettor beware! We are simply speculating that if belief in sports momentum is unwarranted but still influences betting lines, one might find that profitable opportunities emerge as teams go on winning or losing streaks (see Gandar et al., 1988 for more on rationality in NFL betting markets). By the same token, if you believe in the power of sports momentum and it turns out this is mistaken, you might be putting yourself at a disadvantage when you place bets. Our point is not to encourage gambling, but to demonstrate that critical thinking should be heightened as stakes are raised.

If scientific reasoning is to overcome the limitations and biases inherent in System 1 thinking, what are the skills and tools that System 2 needs to use? We review five major components of scientific method that can be used to promote critical thinking, describe how they deal with the shortcomings of System 1, and illustrate them through an examination of sports momentum. In particular, we examine the belief that winning or losing streaks affect the likelihood of future success for NFL teams.

Consider Alternative Explanations to Establish Competing Predictions

System 1 is biased to believe, and to do this it tends to seek only confirmatory evidence that breeds confidence. Alternative ideas may be suppressed, and an important countermeasure is to actively construct alternative hypotheses that make competing predictions. This means asking questions that do not spring to mind by carefully considering the implications of a belief.

Sports momentum clearly entails changes in one's psychological experience following success or failure, but it is less clear that these changes will in turn affect performance. What are the causal mechanisms? It might be that confidence is the key, that a rise or fall in one's feeling of self-efficacy exerts an impact on performance. Those familiar with the Yerkes-Dodson law might speculate that the boost to confidence nudges arousal closer to the optimal level, with the predicted effect being improved performance. Likewise, a fall in confidence could push arousal away from the optimal level and lead to a corresponding drop-off in performance. Plausible as that might sound for behavior unfolding in the moment, it is not obvious how it might be relevant to games played on a weekly basis. Levels of arousal might fluctuate *within* a game, but would this be sustained *between* games? Doesn't the week-long effort of preparing for the next game dissipate the fleeting arousal effects of the earlier wins or losses? Doesn't the immediacy of that next game, as it begins, provide its own motivation to perform and boost arousal? Can athletes even make it into the ranks of professionals if they require an artificial boost to approach their optimal level of arousal?

Perhaps it is not arousal levels that are the causal mechanism for a between-game momentum effect, but an effect on preparation and strategy. Maybe winning or losing streaks prompt a different approach to the next game than a recent history of mixed results? A team on a winning streak might try to continue its success by doubling down on what appears to be working. That might seem reasonable enough, but then again it might be a self-defeating strategy. If a winning team sticks with a style of play that has succeeded, their opponents learn what to expect and can develop ways to mute its impact. Every innovation in strategy is eventually met with an effective response. Even teams that rely heavily on the talent of star players learn that opponents find ways to handle them. Film study and tailoring a game plan to a specific opponent are a huge part of what it takes to win in the NFL, and teams devote a tremendous amount of energy to these tasks. Sticking with what has worked can also be accompanied by complacency, a lack of discipline, and reduced dedication to workouts and practice. None of this is likely to lead to sustained success.

Likewise, it is not hard to imagine that a team on a losing streak might try something new to break the spell of failure, and this may be more effective. Coaches and players alike know that their careers depend heavily on winning, and they are highly motivated to turn things around when they begin to lose. Teams on a losing streak are unlikely to stick with what is not working for long and devising new approaches will keep opponents guessing and make it harder for them to game-plan effectively. Just as the behavioral responses to a winning streak might bring it to an end sooner than expected, the behavioral responses to a losing streak might do the same.

All of this is merely speculation, but the point is that by thinking through the implications of sports momentum one might come to doubt its potency. This requires the effort of critical thinking, not the facile acceptance of uncritical thinking. In the case of sports momentum, the competing predictions would be that it either (a) has the power to influence later performance, in which case one would expect more and longer winning and losing streaks to occur than would be expected by chance, or (b) does not have this power, that it is a cognitive illusion, in which case one would expect outcomes to follow chance-level patterns.

Collect Data as Systematically and Comprehensively as Possible

System 1 relies heavily on introspection, the examination of personal experience and the recollection of instances that appear relevant but that are likely biased due to the search for confirmatory examples. To counter these biases, it is imperative to gather a wider range of pertinent information and to do so as systematically as possible. It is easy to think of NFL teams that have gone on winning or losing streaks, but it is more important to consider all teams' performance to see whether there are in fact more and longer streaks than expected by chance. This clearly exceeds the capacity of System 1 and means turning to archival records and using tools such as spreadsheets to organize and manage these data.

To examine the streakiness of NFL teams, we obtained the specific sequences of wins and losses for every team in every year from 1978, when the NFL moved to a sixteen-game season, through 2016, the latest season that had been completed when these data were collected. This yielded $n = 1,169$ team-seasons, which affords a fairly comprehensive examination of between-game momentum in the NFL.

Establish a Specific, Replicable Protocol

The uncritical thinking of System 1 is free to search for apparent patterns and supportive evidence, which makes it quite easy for questionable beliefs to emerge and to be held with confidence. Critical thinking requires constraints such as clear definitions of relevant constructs and *a priori* plans for research design and data analysis that are expressed in sufficient detail that they can be evaluated on their merits and replicated if desired.

The goal is to remove any wiggle room, to make it harder to ignore the misses and count the hits when tallying results to reach a conclusion.

Before performing any analyses of our NFL game data, we specified the competing predictions made by the sports-momentum and cognitive-illusion hypotheses, decided on the criteria for trimming the sample, and planned all of the analyses that would be performed. Trimming the sample was required because data analysis required that each team-season contained the same number of games, that there was variability within each sequence, and that the outcomes were either a win or loss for each game. Therefore, we removed data for the strike-shortened seasons in 1982 and 1987, for the 2007 New England Patriots (who won all sixteen games) and the 2008 Detroit Lions (who lost all sixteen games), and for any team-season in which there were any tied games. The final sample contains $n = 1,075$ team-seasons. This left most (92 percent) of the original sample intact, and because we removed data using objective criteria that have nothing to do with the pattern of wins and losses within team-seasons, this should not introduce any bias.

As noted earlier, our competing predictions stemmed from the hypotheses that sports momentum either does or does not have the power to influence later performance. We planned a series of five tests, with competing predictions for what we will refer to as a momentum effect and a cognitive illusion. The first test involves the influence of a bye week. The NFL's regular season spans seventeen weeks, but each team plays only sixteen games because they will have one week off. It is widely believed that this bye week, usually scheduled between weeks 4 and 12, will disrupt any momentum a team may have developed, making it harder to continue a winning streak and easier to end a losing streak. The prediction for a momentum effect is that streaks will be disrupted by bye weeks, whereas the prediction for a cognitive illusion is that bye weeks do not affect streaks.

The second test involves the number of runs (streaks) within team-seasons. A team that wins half of its games could do so in a number of ways, such as WWWWWWWWLLLLLLLL or WWLLLWLWWLWWLWLL. The first sequence contains a run of eight wins followed by a run of eight losses, for a total of two runs. The second sequence contains a run of two wins, a run of three losses, a run of one win, and so on, for a total of ten runs. The prediction for a momentum effect is that the number of runs will be smaller than expected by chance, as many of the wins and losses will be clustered together. The prediction for a cognitive illusion is that the number of runs will match chance-level expectations.

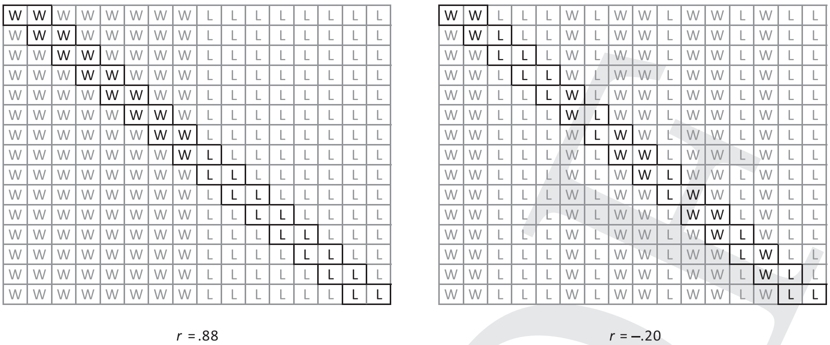

Figure 12.1 Autocorrelations for two illustrative team-seasons
The panel on the left shows a sequence of eight wins followed by eight losses. This sequence is copied across all rows so that each pair of back-to-back games can be visualized separately. These pairings constitute the data points for the autocorrelation analysis, which yields a very large positive correlation because wins tend to follow wins and losses tend to follow losses. The panel on the right shows a sequence in which eight wins are mixed with eight losses in a more haphazard way. The correlation is small and negative, suggesting that wins and losses are distributed fairly randomly throughout the sequence.

The third test involves autocorrelations. This examines the sequential dependence of outcomes, in this case whether wins predict wins and losses predict losses in back-to-back games. Figure 12.1 shows how back-to-back games are paired to form 15 data points within a team-season. A correlation is then calculated by treating each win as a 1 and each loss as a 0. For the first sequence shown above, this analysis yields $r = .88$ because wins strongly predict wins and losses strongly predict losses. For the second sequence, this analysis yields $r = -.20$, suggesting little relationship between wins and losses. The prediction for a momentum effect is that autocorrelations will be larger than expected by chance, whereas the prediction for a cognitive illusion is that autocorrelations will match chance-level expectations.

The fourth test involves predictions. If sports momentum influences performance, this implies that teams' records in more recent game outcomes should exert a larger influence on predictability than game outcomes in the more distant past. As a measure of predictability, we calculated the probability that the teams with the better records won their games at a particular point in the season. Records were calculated using varying numbers of prior weeks' game outcomes. The prediction for a momentum effect is that game outcomes will be more predictable using smaller, more recent samples of teams' wins and losses, whereas the prediction for a cognitive illusion is that game outcomes will be more predictable using larger samples of teams' wins and losses that stretch back farther, as basic statistical theory would suggest.

The fifth and final test involves the frequency of winning and losing streaks. Each team-season can be broken into a series of runs, as shown earlier, and the frequency of such winning or losing streaks of varying lengths across all team-seasons can be tallied. The prediction for a momentum effect is that winning and losing streaks will be longer than expected by chance, whereas the prediction for a cognitive illusion is that the lengths of streaks will be distributed according to chance-level expectations.

Consider the Role of Chance

System 1 is outstanding at identifying patterns and ascribing significance to them. Many patterns identified when thinking uncritically turn out to be illusory, though. To think critically, we must consider the possibility that mere coincidence is at work. Sometimes this can be accomplished by engaging System 2, but often it will require the use of an appropriate statistical analysis to help assess the probability that the pattern in question might occur by chance.

If sports momentum is nothing more than a cognitive illusion, this means that the winning and losing streaks thought to be evidence of momentum effects are no more common than would be expected by chance. In principle, that sounds like a simple comparison to make, but in practice it is not so easy to establish what one might expect by chance. For example, consider what is perhaps the best-known empirical work on sports momentum, the study of the "hot hand" in basketball (Gilovich, Vallone, & Tversky, 1985). Among other things, Gilovich, Vallone, and Tversky examined the probabilities that basketball players made their next shot after having made 1, 2, or 3 previous shots as well as after having missed 1, 2, or 3 previous shots. It turns out that even if the "hot hand" is a cognitive illusion, the expected probabilities are not equal across these conditions. Miller and Sanjurjo (2016) describe a subtle but nontrivial bias in traditional measures of conditional dependence. The solution to this problem is to compare observed results not to a null hypothesis based on one's intuition, but to empirical results for artificial comparison data that are generated using a process that generates truly random sequences.

In our analyses of NFL game data, we generated artificial comparison data by holding constant the number of wins in a team-season and randomly shuffling the order of the wins and losses. Doing this across all 1,075 team-seasons creates one sample of artificial comparison data. Depending on the type of analysis being performed, we generated anywhere from 10 to 1,000 such samples of artificial comparison data to observe the results expected by chance empirically.

Weigh All Available Evidence

System 1 is biased toward seeking support for a belief, of retrieving instances consistent with a potential pattern, and of suppressing doubt. An essential component of critical thinking is to consider all of the available evidence even-handedly. This involves asking whether results converge in support of a conclusion, or whether they are mixed. This does not mean that all sources of information should necessarily be given equal weight, but it does mean that even discrepant findings cannot be ignored outright. As in a meta-analysis, greater weight might be given to larger, better-controlled studies than smaller studies with poorer controls. If the support for a belief consists of anecdotes, even one who rejects this belief needs to account for the anecdotes. For example, one might be able to argue persuasively that the anecdotes are exaggerated, biased and unrepresentative, misunderstandings, or perhaps even fabrications. Other critical thinkers can, in turn, weigh the plausibility of your explanation for any apparently discrepant evidence.

In the case of between-game momentum in the NFL, the evidence begins with the anecdotal observation of winning and losing streaks, such as the beginning of the 2016 season for the Eagles and Vikings, but we expand the evidence base by conducting the five tests outlined earlier. In the first test, we found no evidence of an influence of bye weeks on streaks. Table 12.1 summarizes analyses of the probability of winning after having won (or lost) varying numbers of games across two conditions: the specified sequences of wins (or losses) itself, or this same sequence plus a bye week. None of these analyses yielded statistically significant results, which fails to support a momentum effect and is consistent with a cognitive illusion.

In the second test, we found no fewer runs than expected by chance. Table 12.2 summarizes analyses including all 1,075 team-seasons as well as subsamples based on varying numbers of wins to check the possibility that streakiness might emerge only with sufficiently variable outcomes (e.g., closer to 8 wins than to the extremes of 0 or 16 wins). None of these analyses yielded statistically significant results, which also fails to support a momentum effect and is consistent with a cognitive illusion.

Table 12.1 *Does a bye week affect the probability of winning?*

Sequence (*X*)	*p*(win | *X*)	*n*	*p*(win | *X* + bye)	*n*	*z*	*p*
1 win	.509	3,618	.536	192	−0.73	.465
2 wins	.548	1,605	.560	109	−0.23	.818
3 wins	.550	766	.600	55	−0.73	.468
4 wins	.567	356	.690	29	−1.28	.200
5 wins	.636	162	.667	18	−0.26	.796
6 wins	.654	81	.600	10	0.34	.734
7 wins	.634	41	1.000	3	−1.29	.197
8 wins	.727	22	1.000	1	−61	.544
1+ wins	.537	7,652	.570	423	−1.32	.188
2+ wins	.566	3,591	.597	231	−0.92	.357
3+ wins	.584	1,755	.631	122	−1.02	.307
4+ wins	.613	865	.667	63	−0.85	.395
5+ wins	.638	436	.667	33	−0.34	.738
6+ wins	.647	224	.667	15	−0.15	.879
7+ wins	.632	114	.800	5	−0.77	.443
8+ wins	.673	55	.500	2	0.51	.611
1 loss	.500	3,633	.550	191	−1.33	.184
2 losses	.472	1,601	.427	96	0.86	.390
3 losses	.423	724	.509	55	−1.25	.212
4 losses	.336	351	.500	26	−1.69	.091
5 losses	.372	191	.350	290	0.19	.848
6 losses	.307	101	.333	6	−0.14	.892
7 losses	.246	57	.600	5	−1.70	.089
8 losses	.353	34	.500	4	−0.58	.564
1+ losses	.460	7,640	.493	410	−1.28	.199
2+ losses	.417	3,614	.443	219	−0.75	.455
3+ losses	.372	1,811	.455	123	−1.85	.064
4+ losses	.335	971	.422	64	−1.43	.154
5+ losses	.339	546	.368	38	−0.37	.710
6+ losses	.323	303	.389	18	−0.58	.565
7+ losses	.301	166	.417	12	−0.84	.403
8+ losses	.337	92	.286	7	0.28	.782

A bye week is an off week on a team's schedule during which they do not play a game. The NFL schedule contains seventeen weeks, but each team plays 16 games, with one bye week that's usually scheduled between weeks 4 and 12.

In the third test, we found no larger autocorrelations than expected by chance. Table 12.3 summarizes analyses including all 1,075 team-seasons and, once again, subsamples based on varying numbers of wins. None of these analyses yielded statistically significant results, which once again fails to support a momentum effect and is consistent with a cognitive illusion.

Table 12.2 *Are there fewer runs than expected by chance?*

Team-Seasons	*n*	Observed *M*	Expected *M* (*SD*)	*z*	*p*
All	1,075	7.92	7.86 (2.14)	0.03	.976
1 win	10	2.90	2.88 (0.32)	0.06	.953
2 wins	32	4.47	4.50 (0.77)	−0.04	.972
3 wins	36	6.00	5.88 (1.13)	0.11	.913
4 wins	79	6.96	7.00 (1.42)	−0.03	.976
5 wins	83	7.89	7.87 (1.64)	0.01	.991
6 wins	97	8.60	8.51 (1.80)	0.05	.962
7 wins	122	8.83	8.87 (1.90)	−0.02	.981
8 wins	130	8.90	9.01 (1.93)	−0.06	.956
9 wins	126	8.98	8.87 (1.90)	0.06	.954
10 wins	120	8.88	8.50 (1.81)	0.21	.834
11 wins	95	7.98	7.87 (1.65)	0.06	.949
12 wins	78	7.21	7.00 (1.42)	0.15	.884
13 wins	40	5.75	5.87 (1.12)	−0.11	.912
14 wins	21	4.38	4.50 (0.76)	−0.16	.872
15 wins	6	3.00	2.87 (0.33)	0.38	.704

Each mean (*M*) and standard deviation (*SD*) is for the number of runs per team-season. Expected values were calculated using 1,000 samples of artificial comparison data and treated as population parameters for the *z* tests.

In the fourth test, we found that predictability increased with the number of prior games' outcomes included in the analysis. Figure 12.2 shows that this increasing pattern is observed for each week's games, beginning mid-season (to allow enough prior games for an informative analysis) and continuing through the end of the regular season and ultimately the Super Bowl. The sample sizes become smaller in the postseason because most teams don't make it that far, but the results nonetheless trend in the same direction. The fact that including more games, by reaching farther back into the past, increases predictability relative to smaller samples of more recent games is consistent with a cognitive illusion and not with a momentum effect.

In the fifth and final test, we found no longer or more frequent winning or losing streaks than expected by chance. Figure 12.3 shows that the distributions of both types of streak closely match chance-level expectations. Goodness of fit tests between the observed frequencies and chance-level expected frequencies for these two distributions yielded $\chi 2(9) = 5.13, p = .177$, and $\chi 2(9) = 5.36, p = .198$, respectively. This fails to support a momentum effect and is consistent with a cognitive illusion.

C:/ITOOLS/WMS/CUP-NEW/18876406/WORKINGFOLDER/HALPERN/9781108497152C12.3D 300 [280–308] 10.8.2019 7:15PM

Table 12.3 *Are outcomes for back-to-back games positively correlated?*

Team-Seasons	*n*	Observed *M*	Expected *M* (*SD*)	*z*	*p*
All	1,075	−.08	−.07 (.25)	−0.04	.966
1 win	10	−.06	−.06 (.02)	−0.07	.941
2 wins	32	−.05	−.07 (.20)	0.10	.920
3 wins	36	−.08	−.07 (.23)	−0.05	.963
4 wins	79	−.06	−.07 (.25)	0.01	.994
5 wins	83	−.07	−.07 (.25)	−0.03	.980
6 wins	97	−.08	−.07 (.26)	−0.05	.959
7 wins	122	−.06	−.07 (.26)	0.02	.983
8 wins	130	−.05	−.07 (.26)	0.05	.958
9 wins	126	−.08	−.07 (.26)	−0.05	.959
10 wins	120	−.12	−.07 (.26)	−0.22	.824
11 wins	95	−.08	−.07 (.25)	−0.06	.953
12 wins	78	−.10	−.07 (.25)	−0.14	.890
13 wins	40	−.09	−.07 (.23)	−0.09	.925
14 wins	21	−.02	−.07 (.20)	0.23	.822
15 wins	6	−.07	−.06 (.02)	−0.38	.708

Notes. Each mean (*M*) and standard deviation (*SD*) is for the autocorrelations within each team-season. Expected values were calculated using 1,000 samples of artificial comparison data and treated as population parameters for the *z* tests.

It's easy to see that all five of these test results do in fact converge on a conclusion, namely, that sports momentum appears to be a cognitive illusion. The anecdotal evidence of streaks is not difficult to account for, either, as it is entirely subsumed in the sample of data that we collected and analyzed. The anecdotes were isolated instances of streaks that are bound to occur. A more comprehensive examination of all teams, across all seasons, shows that there is nothing particularly streaky about NFL team performances from week to week. Winning and losing streaks occur about as often, and persist for about as long, as you would expect by chance. Identifying such streaks serves only to describe past performance and appears to hold no predictive value for future performance. In sum, as interesting as it might be to discuss which teams are hot or cold, there is no justification for suggesting that it is harder to face a hot team or easier to face a cold team – at least not beyond what their cumulative win–loss record would otherwise suggest.

Having said this, a final component of critical thinking as scientific reasoning is to acknowledge limitations in the evidence and the extent to which these conclusions can safely be generalized. Despite our best efforts to gather as much pertinent data as we could and to search for evidence of

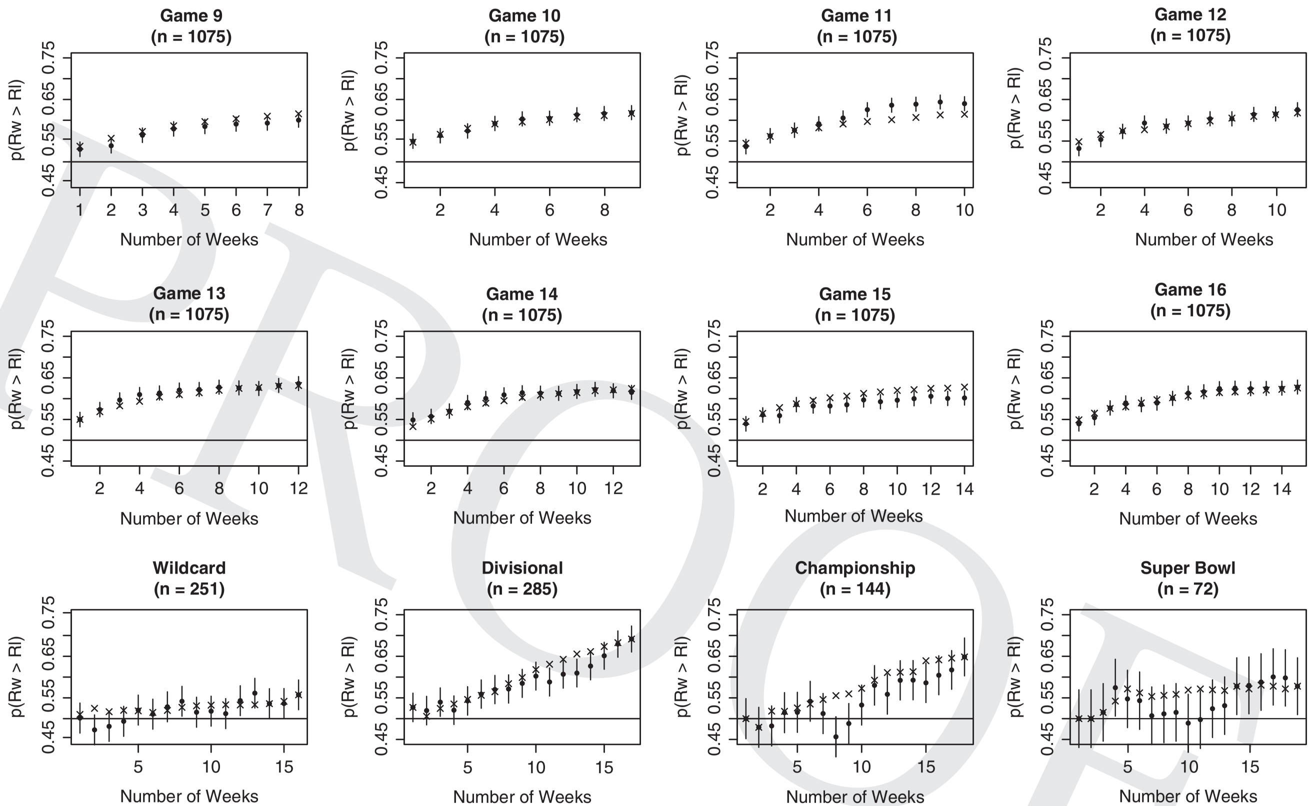

Figure 12.2 Do smaller samples of more recent games predict outcomes better than larger samples that include more distant games? The first graph shows the predictability of Game 9 across all team-seasons using the two teams' records in the previous 1, 2, 3, . . ., 8 games. The *y* axis plots the probability that the team that won the game had a better record than the team that lost the game. Larger values represent greater predictability. Solid circles show the observed values, accompanied by a vertical line representing the standard error. Chance-level expectations were generated using 10 samples of artificial comparison data and plotted as *x*s.

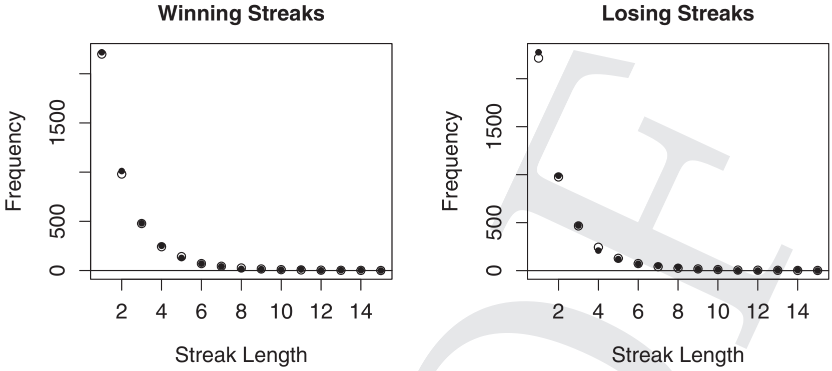

Figure 12.3 Are long streaks observed more often than expected by chance? Solid circles show the total number of streaks of each length across all team-seasons. Chance-level expectations were generated using 10 samples of artificial comparison data and plotted as open circles.

streakiness in many ways, it is possible that a momentum effect escaped detection in our analyses. We compiled data for nearly 40 years of NFL games, so in that regard our sample is both large and representative. However, longer sequences would provide greater statistical power. The sixteen-game NFL schedule provides fairly short sequences within which to search for deviations from chance-level ordering of wins and losses. A small momentum effect may have been missed.

It is also possible that some observed winning or losing streaks might be driven by true momentum effects even as many others are not, in which case the presence of the latter could mask the existence of the former in statistical tests. Even if this is the case, it would imply that the momentum effect must be fairly small, and it remains to be demonstrated whether anyone could successfully differentiate between true momentum and cognitive illusion for particular winning or losing streaks.

Summary and Conclusions

Because it would be mentally exhausting to think critically all of the time, our default mode is uncritical thinking. As Kahneman (2011) notes in his masterpiece *Thinking, Fast and Slow*, the fast thinking of what he calls System 1 is what allows us to process information efficiently. System 1 is adept at pattern recognition and heavily biased toward belief rather than doubt. The extensive literature on judgment and decision-making documents the many cognitive heuristics that generally serve us well but are prone to systematic biases. When the stakes are high enough to warrant the added effort, engaging in critical thinking can help us attain greater accuracy. This requires training System 2 with the tools of scientific reasoning and using them properly on a case-by-case basis.

Critical thinking, undertaken through a series of five components of scientific reasoning, leads us to suggest that sports momentum probably carries little, if any, significance between NFL games. At the same time, we hesitate to generalize to other contexts, such as momentum effects within NFL games or in other sports. For example, it is possible that momentum may operate within NFL games, affecting performance during the course of one particular contest even if it does not carry over into the following week. Two studies have failed to find evidence of such within-game momentum in the NFL (Fry & Shukairy, 2012; Johnson, Stimpson, & Clark, 2012), but further research could alter that conclusion.

As noted earlier, the best-known empirical work on sports momentum began with Gilovich, Vallone, and Tversky's (1985) study of the hot hand in basketball. They concluded that this is a cognitive illusion. Moskowitz and Wertheim's (2011) chapter on momentum discusses related research in many other sports. A meta-analysis of 27 studies and 56 effect sizes drawn from 9 different sports – basketball, golf, baseball, billiards, volleyball, bowling, darts, handball, and soccer, in decreasing order of numbers of effect sizes obtained – corroborates their conclusion (Avugos et al., 2013). Avugos et al. tested various moderator effects as well, finding no practically or statistically significant difference for players vs. teams, between-game vs. within-game analyses, or individual vs. team sports. More recently, Miller and Sanjurjo (2016) challenged the null results for the hot hand in basketball on the grounds that Gilovich, Vallone, and Tversky's measures of streak shooting weren't sufficiently sensitive and at least some of their statistical tests contained biases. Taking both alleged shortcomings into account in their own analyses, Miller and Sanjurjo report evidence in support of the hot hand. It's not clear whether this finding will be replicated by other investigators, that the size of the effect is non-trivial and holds practical significance, or that it generalizes to other sports, but it might be the case that sports momentum can hold some power in certain circumstances.

As this debate continues within and beyond the domain of academic research, the open-minded skepticism of critical thinking will be essential to arrive at accurate conclusions. It would be unwise either to accept or reject all claims for the power of sports momentum, let alone for potential momentum effects beyond sports. For example, Pinker (2011) argues that the *Matthew effect*, which takes its name from Biblical parables and is often summarized as "the rich get richer and the poor get poorer," may apply across societies: "everything seems to go right in some societies and wrong in others" (p. 608). Indeed, as is so often the case, the reality of momentum effects may be nuanced. The uncritical thinking of System 1 will fall far short of the scientific reasoning of a critical thinker in coming to understand these potential subtleties and complexities.

John Ruscio: How Critical Thinking Has Played an Important Role in My Own Professional Career

Like many scientists, I try to be my own harshest critic by looking carefully for any weaknesses in my ideas, methods, or findings. Just as I was leaving graduate school and beginning a job as an assistant professor, I was performing my first study using Paul Meehl's taxometric method. Along with my wife and research collaborator, Ayelet Meron Ruscio, I was trying to determine whether diagnosable depression was a categorical construct (i.e., individuals are either depressed or non-depressed) or a dimensional construct (i.e., individuals vary along a continuum of depressive severity). I ran the taxometric analyses in the conventional manner and interpreted the results in the usual way. The findings seemed clear, but I had some reservations. Though I admired the philosophical foundations of the taxometric method, I found the procedural guidelines vague and the interpretive standards subjective. I put our paper on hold while I spent about six months designing and conducting a simulation study examining taxometric analyses under data conditions similar to ours. The results confirmed my suspicions that the assumptions being made when interpreting taxometric results were in some ways oversimplified. When my wife and I revisited our findings in light of this new evidence, we completely reversed our conclusion. Our report on this research was ultimately published in the *Journal of Abnormal Psychology*. This helped both of our careers at this early stage, and it led to my most productive stream of research over the next twenty years: how to reduce subjectivity in the implementation of taxometric analysis and the interpretation of taxometric results. What I learned along the way has helped me examine, refine, and even develop other statistical methods. Most important have been the lessons that shaped my professional development. I learned firsthand the value of acquiring new skills to complete an investigation as rigorously as possible, of taking the time to explore all doubts rather than rushing to publish initial conclusions, and of thinking critically about my own work through the deliberate, effortful, and slow means of scientific reasoning.

Kevin Brady: How Critical Thinking Has Played an Important Role in My Own Professional Career

Growing up as an avid sports fan, I often took what the television analysts said as fact. After all, they should know more than any casual fan, and would have more insight into what creates wins and losses than someone like me. But as I went through my studies as an undergraduate student majoring in psychology, I began to realize how important critical thinking was and brought this into my day-to-day life. This prompted great interest in studying between-game momentum in the NFL to find if there is a potentially gigantic hole in the way we cover, analyze, and predict NFL games. Momentum is one of those generally-accepted principles which isn't often submitted to critical thinking. The findings of my research project only increased my desire to think through issues critically, and to do so by scientifically questioning evidence. I am still very early in my professional career, but I aim each day to be consciously skeptical and think critically.

Critical Thinking about Critical Thinking Questions

1. Dr. Flurpple is recognized as one of the most accomplished scientists of her generation, with an influential record of scholarly publications and an impressive list of professional honors and awards. At the same time, she holds many superstitious beliefs. For example, Dr. Flurpple always wears her lucky earrings when delivering an important talk, and she refuses to stay on the 13th floor of any hotel because she believes it brings bad luck. How does an understanding of critical thinking as scientific reasoning that engages System 2 help to explain how Dr. Flurpple can be such a successful scientist while thinking uncritically about more mundane issues?
2. Suppose that NFL team A trades for a star player and wins their next several games, team B experiences an injury to a star player and loses their next several games, and a sports analyst says these streaks demonstrate the effects of positive momentum (team A) and negative momentum (team B). What are two plausible alternative explanations for these observed streaks in performance?
3. Cognitive heuristics such as availability, representativeness, or anchoring and insufficient adjustment can lead to systematic biases in judgments or decisions. How is this different from saying that they can lead to random errors? Why is this distinction important?
4. Dr. Glurpple is familiar with the relevant empirical research and finds it hard to understand why so many players, coaches, fans, and analysts

believe in the power of sports momentum. If only people would think more carefully about the alleged streaks in performance that they take as supportive evidence, he argues, they would know better. Aside from recognizing this as an instance of hindsight bias (the feeling that something seems obvious or inevitable after the results are known), what is Dr. Glurpple overlooking with respect to the ability to think critically about a phenomenon like sports momentum? In other words, is it sufficient to "think more carefully" to arrive at a more accurate conclusion?

5. Many academics believe there is a momentum effect in the careers of working scientists. Early success in securing external funding for research and publishing findings in leading scholarly journals is thought to predict later success, and early failure to predict later failure. In what important ways is belief in the power of momentum in scientific careers different from belief in power of sports momentum? How could the former be tested?

Key Terms

Anchoring and insufficient adjustment Cognitive heuristic with which we estimate an unknown quantity by starting with an initial value and moving in the correct direction, but not far enough.

Availability Cognitive heuristic with which we estimate frequency or probability according to the ease with which instances can be recalled.

Cognitive heuristic Mental shortcut that improves efficiency in reaching judgments or decisions.

Representativeness Cognitive heuristic with which we estimate probability or likelihood according to perceived similarity.

Sports momentum The positive or negative change in cognition, affect, physiology, and behavior caused by an event or series of events that affects either the perceptions of the competitors or, perhaps, the quality of performance and the outcome of the competition. Positive momentum is associated with periods of competition, such as a winning streak, in which everything seems to "go right" for the competitors. In contrast, negative momentum is associated with periods, such as a losing streak, when everything seems to "go wrong".

System 1 Effortless, fast, automatic thinking that takes place mostly outside conscious awareness.

System 2 Effortful, slow thinking that requires deliberate attention and control.

REFERENCES

Avugos, S., Köppen, J., Czienskowski, U., Raab, M., & Bar-Eli, M. (2013). The "hot hand" reconsidered: A meta-analytic approach. *Psychology of Sport and Exercise*, *14*, 21–27. DOI:10.1016/j.psychsport.2012.07.005

Fry, M. J., & Shukairy, F. A. (2012). Searching for momentum in the NFL. *Journal of Quantitative Analysis in Sports*, *8*, 1–20. DOI:10.1515/1559-0410.1362

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*, 35–42. DOI:10.1016/j.socec.2010.10.008

Gandar, J., Zuber, R., O'Brien, T., & Russo, B. (1988). Testing rationality in the point spread betting market. *The Journal of Finance*, *43*, 995–1008. DOI:10.1111/j.1540-6261.1988.tb02617.x

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*, 295–314. DOI:10.1016/0010-0285(85)90010-6

Johnson, A. W., Stimpson, A. J., & Clark, T. K. (2012). Turning the tide: Big plays and psychological momentum in the NFL. Sloan Sports Analytics Conference Papers and Proceedings. March 2–3. Online. https://tinyurl.com/6ukyx36

Kent, M. (2006). *Oxford dictionary of sports science and medicine* (3rd ed.). New York: Oxford University Press.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, and Giroux.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–454. DOI:10.1007/978-94-010-2288-0_3

Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York: Norton.

Miller, J. B., & Sanjurjo, S. (2016). A primer and frequently asked questions for surprised by the gamblers and hot hand fallacies? A truth in the law of small numbers (Miller and Sanjurjo 2015). Online. DOI:10.2139/ssrn.2728151

Moskowitz, T., & Wertheim, L. J. (2011). *Scorecasting: The hidden influences behind how sports are played and games are won*. New York: Crown.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review*, *109*, 330–357. DOI:10.1037/0033-295X.109.2.330

Pinker. S. (2011). *The better angels of our nature: Why violence has declined*. New York: Viking.

Sagan, C. (1995). *The demon-haunted world: Science as a candle in the dark*. New York: Random House.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138. DOI:10.1037/h0042769

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. DOI:10.1126/science.185.4157.1124