Descamps, Ambroise, Ke, Changxia, & Page, Lionel
(2022)
How success breeds success.
Quantitative Economics, *13*(1), pp. 355-385.

https://doi.org/10.3982/QE1679

Quantitative Economics 13 (2022), 355–385 1759-7331/20220355

How success breeds success

Ambroise Descamps
Oxera Consulting

Changxia Ke
Schoool of Economics and Finance, Queensland University of Technology

Lionel Page
Economics Discipline Group, University of Technology Sydney

We investigate if, and why, an initial success can trigger a string of successes. Using random variations in success in a real-effort laboratory experiment, we cleanly identify the causal effect of an early success in a competition. We confirm that an early success indeed leads to increased chances of a later success. By alternatively eliminating strategic features of the competition, we turn on and off possible mechanisms driving the effect of an early success. Standard models of dynamic contest predict a strategic effect due to asymmetric incentives between initial winners and losers. Surprisingly, we find no evidence that they can explain the positive effect of winning. Instead, we find that the effect of winning seems driven by an information revelation effect, whereby players update their beliefs about their relative strength after experiencing an initial success.

Keywords. Dynamic contest, momentum, real effort, feedback, confidence, experiment.

JEL classification. C91, D74.

1. Introduction

The famous sociologist Robert Merton coined the term "Matthew effect" (1968) to describe the fact that a success often leads to a string of subsequent successes. In support of this idea, anecdotes abound of high achievers (in business, sports, or academia) describing how one critical success paved the way for what they became (Robertson (2012)). However, whether success breeds success is often debated. While some empirical studies tend to support the existence of a positive effect of past winning on future performance (Malueg and Yates (2010), Gill and Prowse (2012), Mago, Sheremeta, and Yates (2013), Miller, Benjamin, and Sanjurjo (2018), Mago and Sheremeta (2019), Gau-

Ambroise Descamps: ambroise.descamps@oxera.com
Changxia Ke: changxia.ke@qut.edu.au
Lionel Page: lionel.page@uts.edu.au

Available at http://qeconomics.org. https://doi.org/10.3982/QE1679

riot and Page (2019)),[1] others have found either no evidence or the opposite (Ferrall and Smith (1999), Berger and Pope (2011), Fu, Ke, and Tan (2015)).[2]

We investigate this question with an experiment designed to identify both whether there is a causal effect of an initial success (also called *momentum effect*) and, if it does exist, what are the underlying mechanisms. We use the work-horse design of a best-of-three contest with complete information, in which a pair of players compete sequentially in (maximum) three rounds through their performance in a real effort task. The outcome of each round is determined by a stochastic Tullock success function (Tullock (1980)) and a winner arises when one player is the first to win two rounds.

We find clear evidence that players perform better in the second round of a best-of-three contest after winning in the first round, than in the counterfactual situation where they would have lost. We observe an effect on performance both at the extensive margin (time spent on the task) and at the intensive margin (tasks completed per unit of time). When looking at the underlying mechanisms, we find evidence that the effect of winning seems driven by an information revelation effect. Namely, when winning may convey information about their relative strength, the momentum effect exists; whereas it disappears when winning *per se* does not convey this information.

Our paper makes two contributions to the study of the effect of past success on future performance. First, it provides compelling evidence that such an effect exists. One challenge faced by empirical studies investigating momentum is the endogeneity of past performance: past winners may have unobservable characteristics, which are correlated with success in each period. We develop a new empirical strategy to eliminate this concern and identify cleanly the causal effect of winning in a dynamic contest. In a baseline treatment (*Baseline*), we use a stochastic contest success function and leverage the purely random variations in first round outcome it generates. In the first round, conditional on the winning probability, success is entirely exogenous (it is determined by a random draw). We estimate the causal effect of success by matching and comparing winners and losers with identical winning probabilities. This approach delivers a clear result: A momentum effect exists. The first-round winners are substantially more likely to win the second round (20 percentage points increase in winning probability) as a result of both an increase in the productivity and time spent on the task.

The second contribution of our study is to reshape our understanding of the possible mechanisms underlying the momentum. Until now, the debate on the nature of momentum has mostly been articulated on whether it was the result of rational behavior ("strategic") or not ("psychological"). In the economic literature, a momentum typically arises from past success in standard game-theoretic models of contests, assuming

[1]We focus here on the possible effects of success on future performance. Another contributing factor can be the increase in resources and opportunities generated by an earlier success (Van de Rijt et al. (2014)).

[2]A few studies have for instance suggested mechanisms which could lead to a negative momentum because laggards are more motivated to catch up (Tong and Leung (2002), Berger and Pope (2011), Bergerhoff and Vosen (2015)) or because contestants may choke under pressure after an initial success (Paserman and Daniele (2010)). The effect of psychological pressure has however also been assumed to be greater for lagging contestants in some particular circumstances, though the evidence here is debated (Apesteguia and Palacios-Huerta (2010), Kocher, Lenz, and Sutter (2012)).

a complete information setting and competitors with homogeneous ability. This *strategic momentum* emerges due to an asymmetry in incentives between past winners and losers, which generates a positive effect of a contestant's initial success on later performance relative to losers (Harris and Vickers (1987), Konrad and Kovenock (2006), Klumpp and Polborn (2006), Konrad and Kovenock (2009), Fu, Lu, and Pan (2015)). But, outside of standard economic explanations, a momentum arising for psychological reasons and unrelated to strategic reasoning is often mentioned under the term *psychological momentum*. Cohen-Zada, Krumer, and Shtudiner (2017) defined psychological momentum as "the tendency for an outcome to be followed by a similar outcome not caused by any strategic incentive of the players." A substantial literature on momentum in psychology has suggested that past success can increase later performance without any reference to strategic and rational behavior (Bandura (1982), Iso-Ahola and Mobily (1980), Markman and Guenther (2007), Cohen-Zada, Krumer, and Shtudiner (2017), Miller, Benjamin, and Sanjurjo (2018)).

We cast a new light on this debate by looking at the role of information in the emergence of momentum. According to the two main existing explanations of momentum, players' behavior can exhibit momentum due to different sources of information. On the one hand, the strategic momentum emerges from information about the *future*: The expected future rewards for winning the present round. On the other hand, the psychological momentum emerges from information about the *past*: The success or not in the previous round. To investigate the role of information, we toggle on and off different features of the best-of-3 contest which change the informational content of an early success. We replace alternatively the first or the last round of the best-of-three contest with strategically neutral rounds (where the winner is decided by the throw of a die) to eliminate the impact of each source of information. Replacing the first round of the contest with a neutral round, we eliminate the information about a past success while keeping the information about the future reward (treatment *FutureInfo*). The relative positions at the beginning of round 2 are randomly determined, and hence players do not experience a success or a failure leading them to be either ahead or behind. Alternatively, replacing the last round with a neutral round, we eliminate the cost of effort in the last round which is the cause of the asymmetric incentives that results in the standard strategic momentum using backward induction (treatment *PastInfo*). We find no evidence of the *strategic momentum* predicted by traditional game-theoretic models of complete information in the *FutureInfo* treatment, where the first round is replaced. Instead, in contradiction with prevailing game-theoretic explanations, which ignore the role of information, we find clear evidence that past information matters. We observe a momentum in the *PastInfo* treatment, where participants experience success in the first round (with the possible strategic momentum mechanism being switched off).

We investigate further the role of this information about a past success by considering whether the initial success plays a role in changing participants' beliefs about their relative strength. Arguably, real-world contests are never perfectly in complete information since players cannot perfectly know each other's strength (in particular as a player's strength tend to vary over time). However, this does not mean that standard models of contests with complete information are by definition irrelevant, as these models can still be useful if they provide a good enough approximation of real-world contests. For that to be the case, it must be that the effects predicted by the models with complete information exist in real contests and that whatever effects arise from belief updating must be relatively small compared to the effects predicted by the models with complete information. Our treatments *FutureInfo* and *PastInfo* are designed to test this possibility in the best possible scenario for the models with complete information by creating a contest where players are aware the other player is very close to them in terms of characteristics. Nonetheless, small differences in participants' characteristics open the possibility that information revealed through an early success or failure may still play a role in the existence of momentum. To ascertain the possible role of information updating in momentum, we design a fourth treatment, *PastWinUninformative* where all the information about players' performances and winning probabilities in round 1 are revealed to players. Doing so, the winning outcome of the round 1 does not in itself bring any additional information about the players' relative strength. Noticeably, we observe that the momentum disappears in this treatment.

The momentum we observe in the *Baseline* and *PastInfo* treatments seems therefore driven by an *information revelation effect*: an initial success creates a momentum through the informational content it provides to players. Even in contests where contestants are very close in strength, like in our experiment, an initial success may be perceived as containing information about the contestants' relative strength and this information influences their later performance. This result is important for our understanding of the dynamics of momentum in contest. It suggests that traditional models relying on settings in complete information may miss a critically relevant aspect of real-world contests: The fact that players progressively change their beliefs about their relative strength. In the past, deviations from the predictions of these traditional models have been interpreted as evidence in favor of "psychological" momentum. But these deviations may not necessarily be irrational, since players could strategically adapt their behaviors in contests played as games with incomplete information.

We take from these results that understanding behavior in dynamic competitions likely requires a departure from models with complete information to models with incomplete information. Such models may be required to understand the strategic and/or behavioral mechanisms whereby the players' beliefs in their relative strength play a role in their performance. Understanding how self-confidence is shaped by past successes and, in turn, shapes future successes can potentially play an important role in understanding how identical people can end up having very different success paths.

2. A simplified dynamic-contest

Real-world contests are typically embedded in overarching contests whose dynamics are complex to analyze. For instance, being successful in one tender may change a firm's relative position compared to its competitors and influence its strategies in future tenders. Similarly, a successful shot in a sporting match changes the relative scores between teams and, therefore, their optimal strategies. To study the effect of success on future

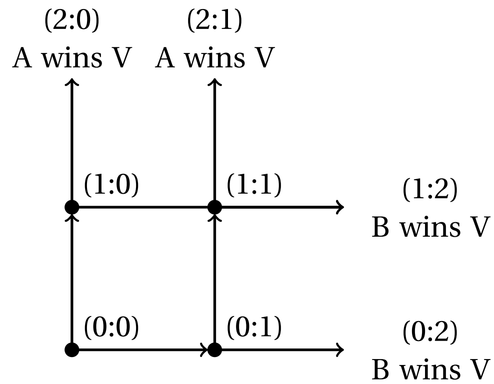

Figure 1. Representation of a best-of-three game.

performance, we design an experiment using the best-of-three game, where the strategic articulation between the different periods is clearly specified. We present here a description of the strategic features of this game. Extended discussions and equilibrium analysis of best-of-N contests can be found in Konrad and Kovenock (2009) and Fu, Lu, and Pan (2015).

Consider a game of complete information with players who are payoff maximizing and homogeneous in terms of ability. They compete over (up to) three rounds. The first player winning two rounds wins a prize V. We denote e_{it} the effort exerted by player i ($i \in \{A; B\}$) in round t ($t \in \{1, 2, 3\}$), and $c(e_{it}) > 0$ the associated cost. The winner of a round is determined according to a contest success function, which assigns a probability of success, depending on a player and his opponent's efforts $(e_{it}; e_{-it})$. Let $p_{it}(e_{it}; e_{-it})$ be this function. Figure 1 represents the structure of such a contest.[3]

The equilibrium strategy of this game is found by backward induction. In round 3, the players face symmetric incentives. Whenever there is a symmetric equilibrium to the game (which is the case for standard contest success functions), both players will have the same expected equilibrium level of effort e_3^*, which induces an expected level of effort cost c_3^*. By symmetry, they have the same expected chance of winning the last round. In equilibrium, the expected payoff of player i when reaching the third round is therefore $v_3 = p_{i3}(e_{i3}, e_{-i3})V - c(e_{i3}) = V/2 - c_3^*$.

In round 2, one of the two players has already won the first round. Without loss of generality, let us assume that it is player A. In case of success in round 2, A gets a value of V. If A is not successful, he gets v_3, the expected payoff from entering the third round. In comparison, B gets a value of v_3 in case of success in the second round and 0 otherwise. The two players' incentives to exert effort only depend on the difference in expected payoffs between winning and losing in round 2 (i.e., the effective prize spread). For A, the incentive is $\delta_A = V - v_3 = V/2 + c_3^*$ and for B it is $\delta_B = v_3 = V/2 - c_3^*$. The incentive

[3]The numbers given in brackets in Figure 1 are the score of player A versus the score of player B at each point in time. Starting from a score of 0:0 at the beginning of the contest, a player can increase his score by one after winning a round. The game ends with one of the four potential outcomes (namely, 2:0, 2:1, 1:2, or 0:2.)

to exert effort is greater for the leading contestant (A), than it is for the lagging one (B): $\delta_A > V/2 > \delta_B$.

This asymmetry in incentives generates a strategic momentum effect whereby the player winning the first round expends more effort in the second round than the player who has lost. Such an asymmetry has been found in a wide range of dynamic contests, leading to a similar strategic momentum effect (Konrad (2009)).

This prediction of a positive effect of success due to a variation in incentives is often opposed to another mechanism whereby *the experience of success* has an effect on performance via its effect on the player's mindset and confidence: "Initial success increases performers' self-confidence and sense of competence and facilitates internal attributions to ability and skills" (Iso-Ahola and Dotson (2014)). This so-called "psychological momentum," is a mechanism favored in the psychology literature.

The notion of self-confidence, which is often seen to play a role in psychology is interesting because it suggests that an initial success may change the belief of the players about their relative strength in the contest. This change of belief is not possible in models with complete information: players already know everything about themselves and the other player and they do not learn anything from winning the first round of the contest. However, real-world contests are never perfectly in complete information. Actual performance depends on skills, and preferences (e.g., preferences for winning), which are never fully observable. The notion of psychological momentum, as used in the literature, therefore possibly blends some behavioral mechanisms, which are not rationalizable, and some mechanisms which may be rationalizable, if the dynamic contest was modeled as a game of incomplete information, where players learn about their relative strength as their observe their initial results.

To further illustrate the various momentum effects, we present in Table 1 a summary of four broad types of momentum according to whether participants have complete or incomplete information and whether the momentum is driven by rational or nonrational behavior.[4] In addition to the traditional "strategic momentum" emerging from rational behavior in games of complete information, other types of momentum could potentially exist. A nonrational momentum in contests with complete information could

Table 1. The possible types of momentum, with examples of possible mechanisms.

	Rational behavior	Nonrational behavior
Contests with complete information	**Rational behavior in complete information** Higher incentives, higher effort	**Nonrational behavior in complete information** Positive mindset, higher effort and greater efficacy
Contests with incomplete information	**Rational behavior in incomplete information** Higher self-confidence, higher effort	**Nonrational behavior in incomplete information** Higher self-confidence, higher effort and greater efficacy

[4]We thank a reviewer for suggesting these distinctions between four types of momentum.

emerge from psychological effects, which are unrelated to self-beliefs. For instance, winning the first round may lead the player to feel good or to have a positive mindset, which makes the player better at performing in the second round. A nonrational momentum in games with incomplete information could also exist due to changes in behavior following changes in self-belief. For instance, a player may feel less stressed and cooler headed when becoming more confident about his relative strength. Finally, a rational momentum could exist in games with incomplete information whereby players learn that they are better and increase their effort level as a consequence. Models of dynamic contests with incomplete information are complex and only a few studies (Konrad (2009), Münster (2009), Ederer (2010), Kubitz (2015), Miklós-Thal and Ullrich (2016)) have provided theoretical results on such models. One of the key challenges is that the incompleteness of information generates signaling motives. Players do not just decide their effort level to win but also to strategically change the beliefs of the other contestant. Nonetheless, we know from the models with complete information that stronger players can have the incentive to expand more effort than weaker players (Konrad and Kovenock (2009)). It is therefore not impossible that players, under incomplete information, learning that they are better after the first round would (rationally) expend more effort in the second round. The term "psychological momentum," which has been used to describe any momentum departing from the standard "strategic momentum" in complete information may actually have been used to describe behavior, which would be rational in games with incomplete information.

3. Experimental design and implementation

We design a best-of-3 contest experiment with the aim to have the best chance of observing the two types of momentum presented in Table 1 (i.e., the rational ("strategic") versus nonrational ("psychological") behavior in contests with complete information). Relative to field studies (Malueg and Yates (2010), Gauriot and Page (2019)), we create a setting as close as possible from a contest with complete information like those investigated by game-theoretic models: players are nearly homogeneous in ability and this fact is common-knowledge among players. We also include a monetary opportunity cost of time spent on the task to ensure that effort is costly enough. We expect these features to give the best chances to generate the type of strategic momentum predicted in game theoretic model in complete information, if it exists. Relative to prior experimental studies (Mago, Sheremeta, and Yates (2013), Mago, Shakun, and Razzolini (2019)), our design innovates by using a real-effort task. We expect this feature to increase the external validity of the design in giving the best chances to generate a "psychological momentum" driven by the fact of experiencing a success after a real performance.[5]

The effort task we adopted is inspired by Huck, Szech, and Wenner (2015): Participants observe on their computer screen a string of 20 characters (numbers, lower and

[5]In chosen effort experiments (Mago, Sheremeta, and Yates (2013), Mago, Shakun, and Razzolini (2019)), players simply indicate the "effort" they are willing to expend and pay a (monetary) price for it. In this setting, prior studies have not found evidence of a "psychological" momentum. Chosen-effort experiments present the advantage of better controlling the cost faced by participants, but they base the study of momentum on a contest without a real performance.

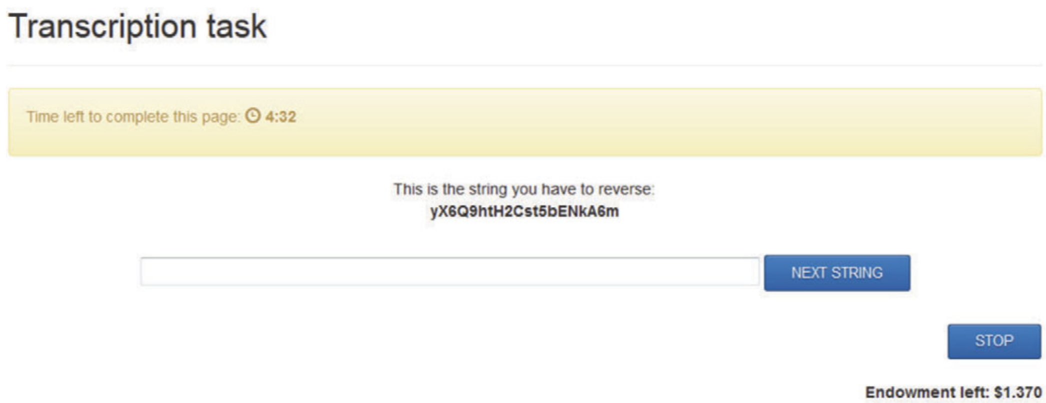

FIGURE 2. A screen-shot of the effort task faced by participants.

upper cases letters), and have to type backwards from the last character in a text-box below it.[6] Each time a string is correctly typed in a reverse order, a new one appears on the screen (see Figure 2 for an experimental screen shot).

The opportunity cost of time is implemented by including a "STOP" button at the bottom-right of the task screen, so that participants can quit working at any time during the competition. Each round of the contest lasts for 10 minutes and participants are endowed with \$3 in total at the beginning. For each second spent on the effort task, participants lose half a cent (\$0.005) of their initial endowment. However, the longer they work on the task, the higher is their performance, and hence the more likely they will win a prize of \$20. Introducing the opportunity cost of time, like in (Gächter, Huang, and Sefton (2016)) and (Erkal, Gangadharan, and Han Koh (2018)), avoids the problem of participants' inelastic response to incentives in real-effort experiments, found by previous literature (Araujo et al. (2016)).

To test for the existence of the strategic momentum effect, we pair players by their abilities and inform them about the pairing to establish common knowledge about the ability difference. It ensures that the contest environment is as close as possible to the game of complete information. Prior to the best-of-three contest, participants enter an initial round where their performances at the task are measured in order for them to be paired by strength in the best-of-three contest. We call this part of the experiment the *evaluation stage*. A player's success in one round of the best-of-three contest is determined stochastically by the Tullock lottery success function, which is most widely used in the literature (Dechenaux, Kovenock, and Sheremeta (2015)). With such a function, the winning probability of player i conditional on his own and opponent's effort (e_{it} and e_{-it}, respectively) in a round t is

$$p_{it} = \frac{e_{it}}{e_{it} + e_{-it}}. \tag{1}$$

[6] To avoid confusion, characters that are too similar were not included, such as capital "O" and zeros.

The lottery function introduces controlled randomness to the outcome of a round, which makes it possible to cleanly identify the causal effect of winning in round 1 on the performance in round 2.

3.1 *Treatment design and testing hypotheses*

In the *Baseline* treatment, we implement a standard best-of-three contest after the evaluation stage. The timing of the *Baseline* treatment is depicted in the top-left corner of Figure 3. After each round of the best-of-three contest, participants are informed of the winner of that round, who is randomly drawn by the computer according to their winning probabilities. Apart from this final outcome, players are not given any additional information, neither the actual number of tasks completed nor the winning probability of each player. In the real world, information is often limited to outcomes with actual levels of effort being unobserved. Our baseline experimental set up reflects such a situation.[7] The *Baseline* treatment is designed to test whether there is any causal effect of winning at all. First, we define the causal effect of winning as follows.

Definition 1 (Momentum effect). In a dynamic contest, there is a positive (negative) momentum effect of early winning on later performance if a player displays a higher (lower) performance after an initial win than in the counterfactual situation where he/she would have lost.

Given this definition of the effect of winning, our first hypothesis is the following.

Hypothesis 1 (Positive momentum). *A positive effect of winning will be observed in the Baseline treatment.*

This hypothesis reflects the widely held idea that winning can have a positive effect on subsequent performance. However, if an effect is observed in the *Baseline* treatment, it is impossible to tell whether it is because players (who have anticipated the later stages of the game) form different expected gains from effort using backward induction or whether they mainly react to the fact of having won in the first round. Both explanations have the same prediction: the first-round winner performs better in the second round relative to his/her matched counterfactual.

The second treatment (*FutureInfo*) is designed to isolate the effect of information about the future stages of the contest. We place players randomly in either a leading or trailing position in round 2. To do so, we assign subjects to either odd or even numbers at the start of the contest, and a computer-simulated die determines the winner in round 1. As the first round is random, it does not provide successful players with the experience of winning as a result of their own performance in round 1. Observing a causal effect of winning in *FutureInfo* would reveal that this effect cannot be explained primarily by

[7] Furthermore, in a game of complete information with homogeneous players, players know their and their opponents' types and observing their opponent level of effort/performance does not have any strategic value in equilibrium.

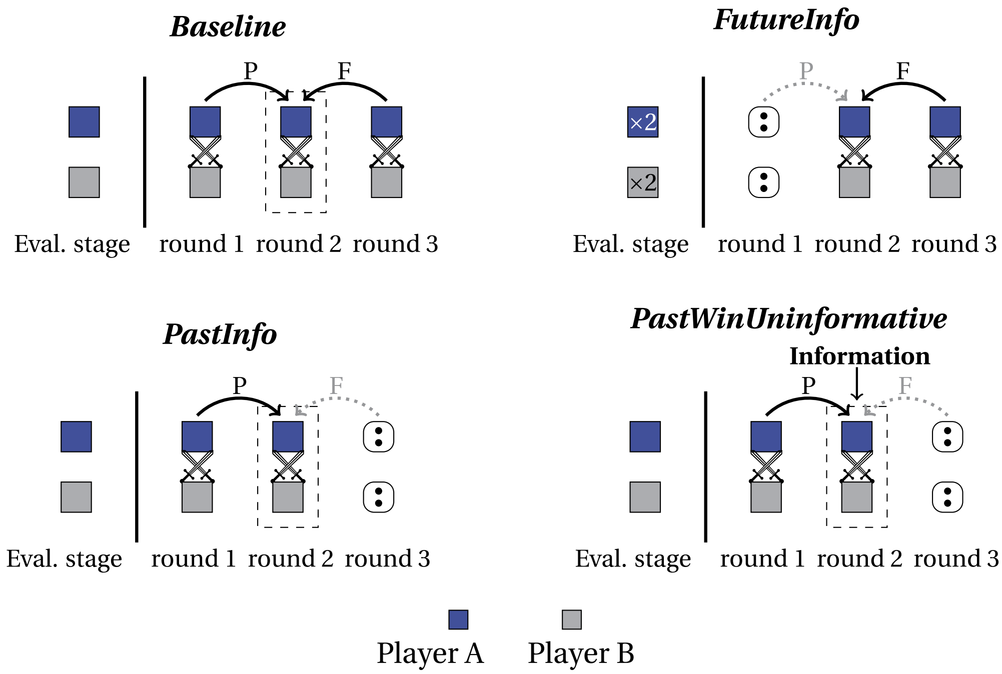

FIGURE 3. Representation of the different experimental treatments. "P" stands for past information, and "F" stands for future information.

the information about a prior success. Instead, this momentum would be compatible with game-theoretic models of contest, where the mere asymmetry of players' positions influence their' strategic effort decisions.[8]

HYPOTHESIS 2 (Effect of future information). *There is a positive effect of winning in FutureInfo treatment, even though the leading players do not experience an initial success that is linked to their performance.*

Note that participants do not exert effort in round 1 in *FutureInfo* treatment. To avoid generating differences in fatigue in round 2 across different treatments, participants play two initial rounds in the evaluation stage before the best-of-three contest in *FutureInfo*. The pairing is similar to other treatments: only the first initial round is used as an evaluation stage to pair participants. The top right panel of Figure 3 shows the timing of this treatment.

The third treatment (*PastInfo*) is designed to identify the effect of past information by turning the possible effect of future information off. As mentioned in Section 2, the expected cost of effort to be expended in the third round causes an asymmetry in incentives in round 2 between round 1 winners and losers. If no effort is to be exerted in

[8]Note that it does not exclude every possible type of "psychological momentum." It is possible that the simple fact of being ahead has a positive effect of mindset even if it did not bring any new information to the players (see Table 1).

the last round, the expected value of going to the last round becomes symmetric ($V/2$ for both players). As a consequence, the round-1 winner's incentive in round 2 (i.e., $V-V/2$) is the same as round-1 loser's incentive in round 2 (i.e., $V/2-0$). We eliminate the cost of effort, and hence the asymmetry of incentives in round 2 by selecting the winner of round 3 randomly, using the same dice procedure as in *FutureInfo*. The bottom left corner of Figure 3 shows the timing of the *PastInfo* treatment. In this scenario, if a player's performance is higher after winning (versus losing) in round 1, it cannot be attributed to standard strategic considerations.[9]

Hypothesis 3 (Effect of past information). *There is a positive effect of winning in PastInfo treatment, even though the usual asymmetry of incentives* (*from backward induction*) *between winners and losers is absent.*

Finally, the fourth treatment (*PastWinUninformative*) extends the third treatment by turning off one possible effect of past information. A natural interpretation of the psychological effect is that players interpret success as giving them information about their relative strength compared to their opponent. We designed our experiment carefully to have matched players in terms of ability. However, it is not possible, neither in experiments, nor in real settings, to eliminate all differences between players. After the first round, players could therefore update their belief about who is the strongest among the two players. To eliminate the possible effect of success on the players' confidence in their relative strength, we modify the *PastInfo* treatment and display all the information about the round 1 performances (including time spent on the task, task completed, winning probabilities of both players) on the computer screen where the winner is revealed at the end of each round.[10] By giving players all this information when round-1's winning outcome is revealed, we remove all the informational content of the outcome. Participants with a winning probability p have the same information about their past performance. Therefore, winners and losers who are matched based on having the same winning probability in round 1 do not differ in their beliefs about their relative strength. Hence, there should be no information-driven effect of success in the *PastWinUninformative* treatment.

Hypothesis 4 (Effect of past information when success is not informative). *Winning per se should not have a causal effect on performance in PastWinUninformative.*

If we observe a momentum effect both in *PastInfo* and *PastWinUninformative*, it would suggest that a psychological momentum exists and that it is not due to an effect

[9]Note that the game structures of the *Baseline* and *PastInfo* are different. Players expect to play two more rounds after a win in *Baseline* and only one more round in *PastInfo*. Therefore, we do not expect an effect of past information to be necessarily equal in the *Baseline* and in *PastInfo*. But if past information has an effect, we would expect to observe an effect of winning in *PastInfo*.

[10]Note that the fact this information will be revealed is not announced in the experimental instructions they receive at the beginning of the contest stage, which means this information comes as a surprise to the participants at the end of round 1. This design ensures that participants have the same information/expectation in both *PastInfo* and *PastWinUninformative* treatments when they compete in round 1.

on players' beliefs about their relative strength. Instead, it could be due to an effect on other aspects of the players' mindset. For instance, successful players may feel elated from their initial success and this feeling could drive a greater performance. However, if we observe a momentum effect in *PastInfo* but not in *PastWinUninformative*, it would suggest that a psychological effect that only works through positive mindset and great efficacy does not exist and the change in psychology must go through confidence changes via information updating.

3.2 *Two ability-pairing conditions*

Ability pairing is crucial in our design to form a common-knowledge of equal ability within a pair, mimicking the complete information setting of game-theoretic models. We implemented our full experimental design in the laboratory using two pairing conditions to ensure our results are robust to how participants are paired.

We first run the experiment with a pairing of participants based on their raw performance. To do so, we incentivize participants to try their best by using a piece-rate payment scheme in the initial round. This round lasts 10 minutes and each transcription task correctly completed is rewarded with 20 cents ($0.2). No timeout "STOP" button is available. The aim of this piece-rate round is to estimate the strength of the players as their raw skill at the task. Pairing players on their raw skills follows closely the assumption of game-theoretic models of contest with homogeneous players.

In practice, however, we could be concerned that players may differ along other dimensions than ability. Players' performance in the contest could potentially differ significantly because of different propensities to use the "STOP" button (e.g., if they differ in "grit"). In that case, pairing players only on their raw ability would not take that heterogeneity into account. To control for this possibility, we also implement another pairing that is based on participants' performance in a single contest in the initial round. The single contest (with a prize value of $7) is resolved in the same way as each round in the best-of-three contest, and a "STOP" button is available. The aim of this one-round contest is to measure the strength of the player using his performance in a single contest environment, which is similar to one round of the dynamic contest. This performance derives not only from the player's raw skill at the task but also from his propensity to stop earlier in a contest. We label these different pairing conditions as "Raw-performance pairing" and "Contest-performance pairing," respectively. Overall our results are very similar across these two conditions.

Participants are then paired based on principles which are common-knowledge to them. In all experimental sessions, *after the evaluation stage and before the beginning of the best-of-three contest*, participants are informed that they will be placed in pairs based on their performance in the evaluation stage. The best performer in the session is paired with the second-best performer, the third is paired with the fourth, and so on. This is the only information they receive about them and their opponent. The information about the contest and the pairing is given only after the evaluation stage to ensure that participants do not play strategically in the evaluation stage.[11]

[11] We follow here the approach of Fu, Ke, and Tan (2015).

3.3 *Data collection*

We ran experimental sessions in two Australian universities of comparable size and student demographics. In each university, a session lasted around 75 minutes, including the time for instructions and payments and the same experimenter ran all the sessions. The experiment was programmed in oTree (Chen, Schonger, and Wickens (2016)). Participants were recruited from various faculties using ORSEE (Greiner (2015)).

The sessions with raw-performance pairing condition as the evaluation stage were run at the Queensland University of Technology (QUT) in 2017: 198 students took part in 14 sessions gathering between 10 and 18 participants each. The mean age of participants was 21, 54.3% were males. The number of participants per treatment was $N = 50$ in *Baseline*, $N = 46$ in *FutureInfo*, $N = 48$ in *PastInfo*, and $N = 54$ in *PastWinUninformative*. The average payment is composed of a \$5 show-up fee, and a variable earning of \$15 on average, ranging from \$0 to \$34.

The sessions with contest-performance pairing condition as an evaluation stage were run at the University of Technology Sydney (UTS) in 2019: 240 students took part in 8 sessions gathering between 22 and 40 students each.[12] The mean age of participants was 24, 56.7% were males. The number of participants per treatment was $N = 66$ in *Baseline*, $N = 58$ in *FutureInfo*, $N = 58$ in *PastInfo*, and $N = 58$ in *PastWinUninformative*. The average payment is composed of a \$10 show-up fee, and a variable earning of \$19 on average, ranging from \$1 to \$40.[13]

The participants received instructions about the initial round(s) both in written form and in PowerPoint slides presentation to ensure common-knowledge before they start the evaluation stage. Afterwards, participants were given a separate set of instructions for the best-of-three contest to read, followed by a short summary presentation. A few control questions were then displayed on their screens, testing their understanding of the game before the game started. At the end of the experiment, participants were asked a few demographic questions. We collected participants' beliefs about their chance to win the next round before each round. We use a simple non-incentivized question: "According to you, how likely are you to win the next round (in %)?".[14] In the sessions ran at UTS, we added several questions: one question about their beliefs in being the best performer within their pair in the next round and four questions to ascertain the players' ability to reason with backward induction, using games designed by Grabiszewski and Horenstein (2019).

[12]We increased the size of the session here on purpose to improve the quality of the pairing of participants by ability.

[13]The difference in average variable earnings between the UTS and QUT data is largely driven by the four incentivized backward-induction games added to the post-experimental questionnaire in UTS sessions. Details about these games will be provided in Section 6.1. We followed the different standards of show-up fee used by each lab which potentially reflects the higher opportunity cost of showing up on UTS campus due to longer commuting time.

[14]It has been shown that simply asking participants their subjective beliefs works well as an elicitation method (Hollard, Massoni, and Vergnaud (2016)).

4. Estimation strategy

We aim to estimate the effect of winning the first round on the performance in the second round. For simplicity, and in line with Section 2's game-theoretic framework, we use "effort" as the main driver of performance in the discussion about the identification strategy. This framework easily generalizes to situations where performance can be affected by other factors, such as the effectiveness of the effort, preference for competing, and risk attitude. To model the effort (e_{it}) in each round t, let us consider the following model:

$$\begin{cases} e_{i1} = \alpha + \delta_1 + u_i + \varepsilon_{i1}, \\ e_{i2} = \alpha + \beta \text{win}_{i1} + \delta_2 + u_i + \varepsilon_{i2}. \end{cases} \tag{2}$$

The variable win_{i1} in (2) is a dummy taking the value 1 if individual i won in round 1, and 0 otherwise. The intercept δ_t is a round specific element which accounts for learning or exhaustion as the participants move through the contest. The term u_i is an individual effect, which accounts for heterogeneity, such as individual differences in ability or preference for competition. Finally, ε_{it}, is a round and individual specific disturbance, which captures residual variations in effort in a given round for a given individual.

The model (2) has, de facto, a dynamic panel data structure since win_{i1} is a function of e_{i1}. As a consequence, usual estimation procedures will deliver biased estimates. It is easy to see that estimating the effect of win_{i1} by ordinary least squares (OLS) suffers from an endogeneity problem. The individual effect (u_i) impacts effort in the first round (e_{i1}), which, in turn, affects individual i's winning chances (win_{i1}). Individuals who exert more effort than their opponent in each round are more likely to win round 1 and round 2. It creates a spurious correlation between the outcome of the two rounds.

Trying to control for this unobserved heterogeneity using a fixed-effects regression also gives biased estimates. The so-called Nickell bias arises from the fact that the fixed effects absorb part of the noises ε_{i2} (Nickell (1981)). It creates an attenuation bias, which can be very large when the panel dimension is short as in our case.

Taking the first difference of equation (2) to try to directly eliminate the unobserved heterogeneity does not work either. The new estimated equation becomes

$$\Delta e_i = \beta \text{win}_{i1} + \Delta\delta + \Delta\varepsilon_i \tag{3}$$

with $\Delta e_i = e_{i2} - e_{i1}$, $\Delta\delta = \delta_2 - \delta_1$, $\Delta\varepsilon_i = \varepsilon_{i2} - \varepsilon_{i1}$. The individual heterogeneity u_i is netted out of the estimation, solving the endogeneity problem of equation (2). However, a different endogeneity problem appears. Random variations of ε_{i1} in round 1 are positively correlated with the winning probability. As a consequence, there is a negative correlation between $\Delta\varepsilon_i$ and win_{i1} and the exogeneity assumption is violated for win_{i1}. We can expect a negative bias in $\widehat{\beta}$ due to a *regression toward the mean*. A win in round 1 partially signals a likely high ε_{i1}, meaning that the effort in round 2 is not likely to be as high in round 2 due to a lower ε_{i2}.

One solution, proposed by Gill and Prowse (2014), is to use the effort of a contestant's opponent as an instrumental variable. The opponent's effort is not correlated with the contestant's effort choice and directly affects his winning chances. However, in our experiment participants are paired with each other by ability in order to closely match the hypothesis of homogeneous ability from the game-theoretic model. Effort levels are therefore highly correlated within a pair, thereby making the instrument invalid.[15]

We develop here a novel approach. Since we use a stochastic contest success function, the result of each round is partly random and the probability of winning of each player is perfectly determined. We therefore use this probability to match winners and losers with similar winning probabilities.[16] Conditional on the winning probability the outcome of the round is purely random: it is the result of a random draw from the computer. We can therefore identify the causal effect of winning using this random variation in winning.[17] To do so, we match winners and losers who have similar ex ante winning probabilities.[18] We use a local linear regression matching, which compares each winner to a weighted average of losers with similar probabilities (Heckman, Ichimura, and Todd (1998)). More weights are given to counterfactual observations with closer matching probability.[19] We use a bandwidth of $h = 0.025$ in our estimations. For example, a round 1 winner with a 50% chance of winning would be compared to round 1 losers with a minimum probability of winning of 47.5% and a maximum one of 52.5%.[20]

As required in matching estimations, we implement a common support restriction suggested by Smith and Todd (2005). We keep in our sample only the set of observations where the empirical distributions of the winning chances of the winners and losers overlap. This restriction ensures that all players who won with a probability p can be compared with players who lost with the very similar probabilities (and vice versa).[21] While bootstrapping fails for nearest neighbor matching, it provides reliable standard errors in the local linear regression case (Abadie and Imbens (2008)).

In the *FutureInfo* treatment, where participants are randomly allocated to a win or loss outcome in round 1 with a 50% chance, the matching procedure is a degenerate case whereby each winner is compared to every loser. It is therefore equivalent to simple OLS.

[15] Let's consider player i and his opponent $-i$. The performance e_{-i1} of the opponent $-i$ in the first round of the contest is likely correlated with his performance in the evaluation stage e_{-i0} (due to unobserved heterogeneity u_{-i}). Given the pairing, this performance is itself correlated with the performance of the player i in the evaluation stage, e_{i0}, and, therefore, it is also correlated with the performance e_{i1} of the player i in the first round of the contest (due to unobserved heterogeneity u_i).

[16] This approach is similar to a propensity score matching procedure (Todd (2010)), but contrary to most propensity score matching applications we *perfectly* know how the probability of an observation being in one or the other conditions. It is precisely determined by a function taking the observed performance of players as input. Our matching approach is therefore not threatened by the possible residual effect of unobservable characteristics on the outcome (this concern typically arises when propensity score matching is used with field data).

[17] Consider, for example, players who had 60% chances of winning the round 1 based on their performance (and the performance of their opponent) in that round. Some will be successful and some will not, due to the computer's random draw. The situations where players have a 60% chance of winning round 1 may be heterogeneous, but the win/loss result will be randomly allocated among these players.

[18] See Appendix B for a formal description of the matching approach.

[19] As shown by Fan (1992), local linear regression are also a form of local weighted averaging.

[20] We show that our results are robust to other choices of bandwidth in the online Appendix C.3.

[21] We report the standard errors constructed by standard bootstrap in the main text. As a robustness check, we also provide the standard errors using bootstraps at the pair-level in the Appendix of the Online Supplementary Material (Descamps, Ke, and Page (2022)) to account for the nonindependence of observations within pairs. Our main results are stable to the way we bootstrap standard errors.

5. Main results

5.1 *Overview of the data*

Our random allocation of participants across treatment did not lead to noticeable differences in their characteristics.[22] Table 2 shows the number of strings completed correctly overall and by treatment, for each pairing condition and when pooling observations from both pairing conditions.[23] The performance levels vary slightly in the evaluation stage between two pairing conditions, which were run on different locations. When pooling observations across pairing conditions, participants completed on average around 16 strings in the initial round and between 17 to 22 strings in the best-of-three contests. There is no significant difference in average performance in the evaluation stage across treatments whether it is compared within each condition (p-value $=$ 0.521 and 0.983) or when all observations are pooled (p-value $= 0.802$).

After the evaluation stage, participants are paired by performance. The resulting within-pair differences of performances are fairly small. The average difference in performance within a pair is 2 when the evaluation stage is a piece-rate round and 1.15 when it is a one-round contest. It is also reassuring that there is no significant difference in average performance in the first round of the best-of-three contests across all treatments (p-value $= 0.824$).[24]

Our main variables of interest are the measures of performance in round 2 as a function of the outcome of round 1. Figure 4 summarizes the performance (as the number of strings correctly reversed) and time spent on the tasks in round 2 and the change of performance and time spent from round 1 to round 2, by treatment and by the winning outcome in round 1. The patterns emerging are in line with the expected direction of the biases arising from the direct comparisons of performance, due to the endogeneity issue discussed in Section 4.

When looking at absolute performance and time spent (left panels), winners in round 1 seem to complete more strings and spend more time in round 2. This comparison is similar to estimating equation (2). This observed difference may reflect a selection bias: Winners in round 1 may just happen to be better at the task.[25] On the contrary, when looking at the changes in performance and time spent between round 1 and round 2 (right panels), we do not observe any clear difference between winners and losers of the round 1. This comparison is similar to estimating equation (3). The absence

[22] Table 11 in the Appendix in the Online Supplementary Material provide detailed statistics on the participants' demographics per treatment.

[23] Tables 7 and 8 in the Appendix provide more detailed summary statistics.

[24] Note that to control for learning or fatigue (as is explained in Section 3.1), participants in the *FutureInfo* treatment were randomly rematched after the first round of the evaluation stage and were requested to compete in another single-round contest. The repetition came as a surprise to them because we wanted to keep everything comparable in the initial round, which we later use for ability pairing. It is also reassuring to note that the average performance in the second single-round contest in the *FutureInfo* treatment is 18.40 for all observations, which is very similar to the average round-1 performance in other treatments where there were only one round in the evaluation stage.

[25] While such a concern should be alleviated with our pairing of contestants by abilities, residual differences can exist.

TABLE 2. Comparison of performances overall and by treatment and pairing conditions. Only the first round performance in the evaluation stage is presented for the *FutureInfo* treatment in order to have a comparable situation to other treatments. Note that, by design, the overall number of observations is the sum of observations in *baseline*, *PastInfo*, and *PastWinUninformative* (shorten as *PastWinUn*) in round 1 and the sum of observations in *baseline* and *FutureInfo* in round 3.

		Number of strings					Kruskal–Wallis test
		Overall	*Baseline*	*FutureInfo*	*PastInfo*	*PastWinUn*	
				All observations			
Eval. stage	Mean	16.35	16.45	16.44	16.03	16.03	$p = 0.802$
	(sd)	(7.83)	(8.75)	(7.34)	(7.29)	(7.40)	
	N	438	116	104	106	112	
Round 1	Mean	18.68	18.59	–	18.92	18.33	$p = 0.824$
	(sd)	(8.44)	(9.23)	–	(7.53)	(7.52)	
	N	334	116	–	106	112	
Round 2	Mean	18.97	19.16	18.45	18.82	16.77	$p = 0.136$
	(sd)	(8.22)	(9.30)	(7.62)	(7.45)	(8.12)	
	N	438	116	104	106	112	
Round 3	Mean	21.74	21.57	21.93	–	–	$p = 0.583$
	(sd)	(7.04)	(7.18)	(6.96)			
	N	86	44	42			
				Raw-performance pairing			
Eval. stage	Mean	20.12	20.84	20.35	19.92	19.44	$p = 0.521$
	(sd)	(5.79)	(5.38)	(4.80)	(6.25)	(6.53)	
	N	198	50	46	48	54	
Round 1	Mean	19.01	19.96	–	19.64	17.56	$p = 0.114$
	(sd)	(8.21)	(8.17)	–	(7.67)	(8.65)	
	N	152	50	–	48	54	
Round 2	Mean	18.98	20.30	18.63	19.54	17.57	$p = 0.116$
	(sd)	(7.82)	(8.22)	(7.31)	(6.92)	(8.54)	
	N	198	50	46	48	54	
Round 3	Mean	22.38	21.63	23.14	–	–	$p = 0.335$
	(sd)	(6.31)	(5.33)	(7.21)			
	N	44	22	22			
				Contest-performance pairing			
Eval. stage	Mean	13.03	13.12	13.34	12.81	12.84	$p = 0.983$
	(sd)	(7.65)	(9.36)	(7.55)	(6.52)	(6.75)	
	N	240	66	58	58	58	
Round 1	Mean	18.27	17.56	–	18.31	19.05	$p = 0.530$
	(sd)	(8.08)	(9.88)	–	(7.42)	(6.27)	
	N	182	66	–	58	58	
Round 2	Mean	17.73	18.30	18.31	18.22	16.02	$p = 0.548$
	(sd)	(8.49)	(10.02)	(7.91)	(7.87)	(7.71)	
	N	240	66	58	58	58	
Round 3	Mean	21.07	21.50	20.60	–	–	$p = 1.000$
	(sd)	(7.73)	(8.88)	(6.59)			
	N	42	22	20			

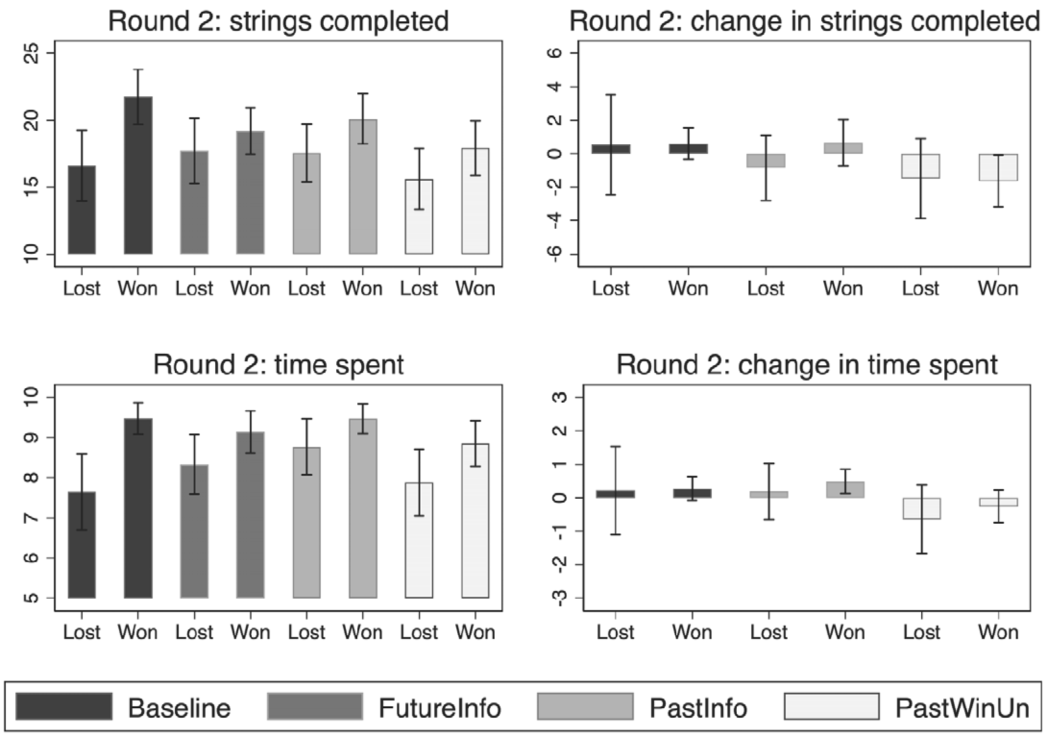

Figure 4. A summary of performance (measured by the number of tasks completed) and effort (measured by time spent on the tasks) in round 2 (left panels) and variation of performance and effort between round 2 and round 1 (right panels). The right panels do not include the *FutureInfo* treatment, since participants do not need to make efforts in the first round there.

of difference in the change of performance may reflect a risk of regression towards the mean: Participants who happen to perform unusually well in round 1 are more likely to win; In round 2, these participants may be more likely to be back to a normal (lower) level of performance. Such a selection pattern could explain why no difference is observed when using first differences.

Our identification strategy is specifically designed to solve the endogeneity issue arising when looking at levels and variations of performance and effort. The probability matching procedure compares winners in round 1 to their losing counterfactual who had very similar probabilities of winning in round 1.[26]

5.2 *Identifying the momentum effects*

Table 3 shows the average treatment effect of winning on different "effort" measures and on participants' winning probabilities. All results refer to the estimation of equation (6) (in the Appendix) where the causal effect of winning is calculated as the difference

[26]This identification strategy can only be used to look at the effect of a success in round 1 on round 2 performance. It cannot, by design, be used in round 3. The matching on round 2 winning probabilities would not control for differences in performance in round 1. See Appendix D for a more detailed discussion.

Table 3. Effect of winning round 1 on "effort" measures and winning probability in round 2, estimated by LLR matching. For estimates with all observations, the matching is done within each pairing condition. The bandwidth for the LLR is set to 0.025 with an epanechnikov kernel weighting function. Standard errors are constructed by standard bootstrap (2000 replications). n is the total number of observations in each treatment and n (supp.) is the total number of observations on the common support. Confidence intervals at 95% are indicated in brackets.

Treatment	Strings completed	Time spent	Productivity	Winning prob.	N	N (supp.)
		All observations				
Baseline	4.01	1.86	0.38	0.20	116	90
	[0.6, 7.43]	[0.49, 3.23]	[0.09, 0.67]	[0.09, 0.3]		
FutureInfo	0.90	0.54	0.01	0.05	104	104
	[−1.5, 3.3]	[−0.44, 1.52]	[−0.13, 0.15]	[−0.02, 0.11]		
PastInfo	2.67	0.63	0.17	0.08	106	100
	[0.49, 4.85]	[−0.21, 1.46]	[0.02, 0.33]	[0.03, 0.13]		
PastWinUninformative	0.33	0.56	−0.01	0.04	112	94
	[−2.84, 3.51]	[−0.72, 1.85]	[−0.21, 0.2]	[−0.04, 0.12]		
		Raw-performance pairing				
Baseline	4.46	1.96	0.65	0.26	50	32
	[−0.34, 9.26]	[−0.08, 3.99]	[0.08, 1.23]	[0.09, 0.42]		
FutureInfo	0.30	0.28	−0.03	0.02	46	46
	[−3.15, 3.76]	[−1.09, 1.65]	[−0.23, 0.18]	[−0.06, 0.1]		
PastInfo	2.93	0.52	0.26	0.08	48	46
	[−1.54, 5.48]	[−1.17, 1.57]	[−0.04, 0.5]	[−0.04, 0.13]		
PastWinUninformative	1.46	1.15	−0.06	0.09	54	38
	[−3.72, 6.65]	[−0.8, 3.09]	[−0.33, 0.2]	[−0.03, 0.2]		
		Contest-performance pairing				
Baseline	3.76	1.81	0.23	0.16	66	58
	[−0.92, 8.45]	[−0.01, 3.63]	[−0.09, 0.55]	[0.03, 0.29]		
FutureInfo	1.38	0.75	0.04	0.06	58	58
	[−1.71, 4.47]	[−0.6, 2.09]	[−0.16, 0.24]	[−0.02, 0.15]		
PastInfo	2.46	0.72	0.10	0.09	58	54
	[−0.4, 5.31]	[−0.34, 1.77]	[−0.06, 0.26]	[0.02, 0.15]		
PastWinUninformative	−0.44	0.17	0.03	0.01	58	56
	[−4.5, 3.63]	[−1.53, 1.87]	[−0.26, 0.32]	[−0.1, 0.12]		

between a player's change of effort from round 1 to round 2, Δe_i, to his counterfactual, $\Delta\widehat{e_i}$.[27] Table 3 presents the result for the two different pairing conditions separately in the top two panels. The results are very similar in both conditions. The last part of Table 3 presents the estimation on the aggregated data.[28] Aggregating the two conditions allows us to gain greater statistical power, and hence we focus on the aggregated results in the following discussion.

[27]For the *FutureInfo* treatment, we use the difference with the evaluation stage: the first of the initial round that was used to pair players. As round 1 success is purely random, there is no risk of regression towards the mean in this specific treatment.

[28]In the aggregated estimations, observations are matched with observations from the same pairing condition.

In *Baseline*, we observe a significant positive momentum on all performance measures. At the extensive margin, we find that winners spent on average more time in round 2 than losers (difference of 1.86min, $p = 0.01$). At the intensive margin, winners display a higher productivity. On average they complete 0.38 more strings per minute ($p = 0.01$). These two joint effects combine into a significantly larger performance for winners who complete 4.01 strings more in round 2 ($p = 0.02$). It means that participants, with similar probability of winning in round 1, diverge in their performance in the second round based on their success in the first round. With this increase in performance, the round 1 winner increases his/her chance of winning in round 2 by 20 percentage points ($p < 0.001$).

Result 1 (Positive momentum). *We observe a positive causal effect of winning on performance in the Baseline treatment.*

The origin of this effect is a priori unclear. It can reflect a strategic momentum due to asymmetric incentives to expend effort after winning (versus losing), or a psychological momentum, or both. Taken on its own, the result from the *Baseline* treatment does not allow us to disentangle the two possible effects. In order to identify the source of the effect, we now turn to the *FutureInfo* and *PastInfo* treatments.

In *FutureInfo*, we do not observe a significant difference in round 2 between the first round winners and losers. Randomly assigned round 1 winners do not spend significantly more time in the second round (diff $= 0.54$, $p = 0.28$), they do not become more productive (diff $= 0.01$, $p = 0.89$) and they do not complete more strings in round 2 (diff $= 0.90$, $p = 0.46$). This absence of effect is found in both pairing conditions.

Result 2 (No effect of future information). *We do not observe an effect of winning when we remove the past information* (*even if the standard strategic momentum is still predicted to be present*).

In *PastInfo*, we find that winning has a significant impact on productivity and the number of strings completed. Winners spent on average 0.63 minutes more on the task, though the difference is not statistically significant ($p = 0.14$) at the conventional level. The productivity of winners is higher by 0.17 tasks per minute ($p = 0.03$). Overall, the performance of round 1 winners is greater by an average of 2.67 strings in round 2 ($p = 0.02$). This greater performance is associated with an increase of winning chances of 8 percentage points ($p < 0.001$). As winners are randomly determined in round 3 in *PastInfo*, the information about the future of the contest is turned off. The presence of a positive effect in *PastInfo* points to the existence of an effect of past information in the contest.

Result 3 (Positive effect of past information). *We observe a positive causal effect of winning on performance in the PastInfo treatment.*

The joint results of these three experimental treatments provide two clear conclusions. First, winning has indeed a positive causal effect on players' performance. Second, the evidence point toward an effect of having experienced a success in the past.[29]

Looking at the role of information updating with the *PastWinUninformative* treatment, we observe that there is no significant effect of winning when success per se does not provide additional information to the player. In the aggregated results, winners did not spend significantly more time than losers in round 2 ($p = 0.39$), they did not complete more strings per minute ($p = 0.94$) and, as a consequence, they did not significantly increase their number of strings completed relative to losers ($p = 0.84$). We can therefore not reject the null hypothesis that there is no momentum effect in *PastWin-Uninformative*, in line with the prediction from Hypothesis 4. This result suggests that the momentum observed in *PastInfo* is driven by the informational content of a win.

Result 4 (No momentum when success is not informative). *We do not find evidence of a causal effect of winning on performance when winning does not give participants any additional information about their relative strength in the contest.*

6. Investigating the mechanisms

We find an absence of the momentum caused by the standard strategic effect from future information and an existence of positive momentum caused by the informational content of past win associated with the experience of a successful performance. In this section, we further investigate and discuss the possible mechanisms underlying these results.

6.1 *Sophistication*

A possible explanation for the absence of a strategic effect of winning could be the participants' cognitive limitations: they may fail to appreciate the strategic aspects of the whole contest. By backward induction, the strategic effect should arise if participants take into account the cost of future effort in round 3. It is this cost which creates an asymmetry in incentives in round 2 between winners and losers.

Failures to fully reason by backward induction have previously been observed (Binmore et al. (2002), Johnson et al. (2002)). We therefore investigate whether it could be the reason that participants fail to display a strategic effect of success.[30] For this purpose, we include four games assessing the degree to which participants are able to do backward induction in the post-experiment questionnaire when we ran sessions for the

[29]The coefficients are slightly different in magnitude in the *Baseline* and the *PastInfo* treatments. These differences are not statistically significant ($p = 0.52$, 0.14, and 0.23 for strings completed, time spent, and productivity, respectively). Note that, as discussed in Section 3.1, the *Baseline* and the *PastInfo* treatments have different game structures. The momentum could potentially have different magnitudes in these two treatments. For instance, a psychological momentum could be smaller in *PastInfo* than in *Baseline* because participants only anticipate one more round of contest in *PastInfo*, versus two in *Baseline*. Therefore, we are primarily interested in the existence or not of an effect in each treatment (not on them being equal).

[30]We thank an anonymous reviewer for this suggestion.

(a) Graphical representation in the game

(b) Extensive-form representation of the game

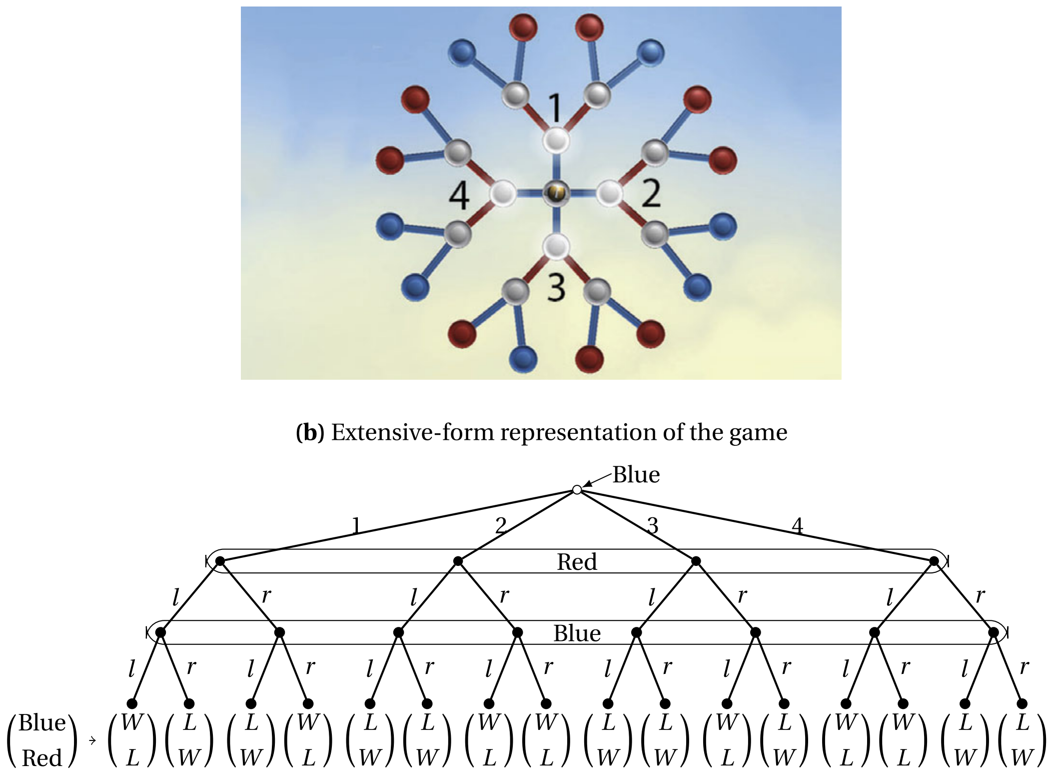

FIGURE 5. An example of the game with two steps of backward induction. Participants were invited to play as the blue player and to indicate which initial choice they would make, starting from the centre to ensure that the outcome of the game is a blue node. After the initial choice, players successively take turn playing left (l) or right (r). Only one possible choice in the first move ensures a win. In this example, only the choice of 1 ensures a win (W) for the blue player.

contest-performance paring condition. These games are taken from Grabiszewski and Horenstein (2019). They are graphical representations of extensive-form games where two players play alternatively. The two players are represented by the colors blue and red, respectively. The end nodes of the game are associated either with a blue or a red color, which indicates the final winner. The blue player is the first mover and his/her goal is to secure the possibility to land on a blue node at the end of the game. Figure 5 shows one of the game in its graphical representation (used in the experiment) and its corresponding extensive form.[31]

Each game starts with four possible moves for the blue player and then the red and blue players play alternately following the node picked by the blue player in the first move, with two possible moves each time afterwards. The player is asked what he/she would do as the first (blue) mover to win the game. Before playing, participants were shown explanations about the game with one simple example. We selected four games

[31] All the games are depicted in Figure 10 in Appendix F.

Table 4. Effect of winning round 1 on “effort” measures and winning probability in round 2, estimated by LLR matching. Observations in the *FutureInfo* treatment under the contest-performance pairing condition are split into two sub-samples by the number of backward-induction games participants did correctly. The bandwidth for the LLR is set to 0.025 with an epanechnikov kernel weighting function. Standard errors constructed by standard bootstrap (2000 replications) and are indicated in brackets.

	FutureInfo treatment			
	Strings completed	Time spent	Productivity	Winning prob.
Numb. games right ≤ 2	1.31	0.79	0.07	0.06
(N = 36)	(2.18)	(0.90)	(0.14)	(0.06)
Numb. games right > 2	0.27	0.26	−0.09	0.03
(N = 22)	(2.35)	(0.76)	(0.20)	(0.07)

where the blue player has to choose between four possible actions as a first move. Only one of them can, by backward induction, ensure a win for the blue player. The games had different degrees of difficulty with participants having to do backward induction over 1, 2, 3, and 4 steps, respectively. Participants were paid one Australian dollar for each correct answer. The results of the participants reflect this increasing difficulty with the proportions of right answer being 95%, 57%, 60%, and 11% in the order of difficulty of the games. On average, participants answered correctly to 2.2 games.

In order to assess whether strategic sophistication and in particular the ability to do backward induction influences the presence of strategic momentum effect in *FutureInfo* treatment, we estimated the causal effect of winning round 1 on two sub-samples: the sample of participants having solved two games or less and the sample of participants having solved more than two games. Table 4 shows the results. We find no evidence that participants who solved more of the backward induction games display a greater effect of success. It suggests that participants did not differ in their strategic behaviour based on their ability to do backward induction.

6.2 *Beliefs and confidence*

In spite of our pairing of players by ability, there may be residual differences. Even if there are no significant differences, the information about a past success could still have an impact if players believe that residual differences in ability exist and update their beliefs accordingly. While our design aims to create a setting as close as possible from common-knowledge of equal ability, we cannot rule out the possibility that players think they differ somewhat from each other. For instance, a player may believe that he could have done better in the initial round and, therefore, is likely paired with a weaker player.

A conjecture that motivated our Hypothesis 4 is that the effect of winning may be driven by participants updating favorably their belief in their relative strength after a success, which further influences their performance or incentive to make effort. To assess this prediction, we asked participants how likely they thought they would win the

Table 5. Effect of winning round 1 on elicited confidence in round 2's winning chances and relative performance, estimated by LLR matching. For estimates with all observations, the matching is done within each condition. The bandwidth for the LLR is set to 0.025 with an epanechnikov kernel weighting function. Standard errors are constructed by standard bootstrap (2000 replications). Confidence intervals at 95% are indicated in brackets.

Treatment	Win. chances	Performance	N
	All observations		
Baseline	20.52		116
	[14.34, 26.69]		
FutureInfo	11.73		104
	[6.14, 17.32]		
PastInfo	16.24		106
	[9.07, 23.42]		
PastWinUninformative	10.08		112
	[1.45, 18.7]		
	Raw-performance pairing		
Baseline	18.58		50
	[9.64, 27.53]		
FutureInfo	9.96		46
	[2.2, 17.71]		
PastInfo	14.76		48
	[3.94, 25.95]		
PastWinUninformative	7.23		54
	[−2.55, 17.01]		
	Contest-performance pairing		
Baseline	21.58	19.47	66
	[13.47, 29.7]	[10.83, 28.12]	
FutureInfo	13.14	8.28	58
	[5.23, 21.05]	[−0.57, 17.12]	
PastInfo	17.51	13.48	58
	[8.32, 26.7]	[5.15, 21.82]	
PastWinUninformative	12.01	−0.24	58
	[−0.42, 24.44]	[−13.54, 13.05]	

next round before each round started (except for the neutral rounds). Given that winning is partly random with the stochastic contest function, we also asked participants how likely they thought their performance would be higher than their opponent in the next round (but this question was only presented in the contest-performance pairing sessions). If the informational content of a win influences confidence, players' beliefs in their chance of winning and in their performance being better should be higher in *Baseline* and *PastInfo* treatments than in the *PastWinUninformative* treatment. Using our matching approach, we can estimate the causal effect of winning on players' confidence in round 2. Table 5 shows the results of these estimations.[32]

[32]See Table 9 and 10 in the Appendix in the Online Supplementary Material for detailed summary statistics of the beliefs elicited in all rounds.

In all treatments, winning leads participants to become significantly more confident when considering chances of winning. The observed effect is the largest in *Baseline* (20.52). This effect is slightly smaller in *PastInfo* (16.24), but becomes significantly smaller in the *PastWinUninformative* treatment (10.08, $p = 0.02$).[33] Interestingly, the effect in *PastWinUninformative* has the similar level as in *FutureInfo* (11.73, $p = 0.69$) where winning is random in the first round and players do not learn anything about their relative strength from winning. The fact that the effect of beliefs is significant in *PastWinUninformative* and *FutureInfo* suggests that the mere fact of winning a random contest in round 1 does impacts participants' beliefs about their chances of winning in round 2.

When we ask participants about their belief of being the better performer in round 2 (rather than about their chances of winning), we observe that the effect of winning becomes close to zero (-0.24) and not significant in *PastWinUninformative* ($p = 0.97$), even though it stays high for participants in the *Baseline* (19.47, $p < 0.01$) and *PastInfo* (13.48, $p = 0.01$). This question has the advantage of excluding any possible effect of winning on the participants' perception about the randomness of the contest (e.g., "feeling lucky" after winning the first round).[34]

To investigate the possible effect of confidence on performance, we first regress the change in effort between the two rounds on the change in confidence, controlling for the winning probability. The results are presented in Table 6. We observe that changes in confidence are positively correlated with an increase in effort in the *Baseline* treatment ($p = 0.04$). The effect just fails to be significant at the conventional 5% level in the *PastInfo* treatment ($p = 0.056$). This indicates that, for a given probability of winning in round 1, participants who experienced a larger increase in confidence also had a larger increase in performance.

These regressions results are supportive of a possible role of the change in confidence in the change of performance. Our setting allows us to look further into this possibility. The win outcome in the round 1 is random conditionally on the winning probability. We can therefore use an initial win as an instrument for confidence to look at the effect of an increase in confidence, induced by an initial success, on performance. The lower part of Table 6 also presents these results. We observe that confidence has a positive effect on performance. The effect is the strongest and significant only in the *Baseline* ($p = 0.01$) and *PastInfo* treatment ($p = 0.04$) where an initial success is potentially informative about the relative strength of the players.

These results do not definitely prove that confidence has a causal effect. A win could have other indirect effects via different variables. Nonetheless, they are in line with the idea that a higher self-confidence may drive a higher performance after a success. In psychology, lack of confidence is often considered to be associated with "self-doubts,"

[33] Similarly, we focus on the pooled data while discussing beliefs on winning chances.

[34] Winners in *FutureInfo* are still marginally more likely to believe that they will *perform* better in round 2 after winning round 1 even though the outcome in this round was determined randomly by the computer (8.28, $p = 0.07$). These results might suggest, after an initial win, participants are not just feeling lucky about their future winning chances but also more confident about their future performance, even though winning was purely random and should not give them any feedback about their relative strength.

Table 6. Regressions of the effect of the change in confidence on the change of effort from round 1 (R1) to round 2 (R2), controlling for each contestant's winning probability in R1. In the IV regressions, the change in confidence is instrumented by whether or not the contestant had won R1. Note that in *FutureInfo* the winning probability in R1 is 50% for every contestants as R1 winning outcome was determined randomly, and hence is omitted. For the same reason, the effort in the second round of the evaluation stage is used as the R1 effort in *FutureInfo*. Confidence interval at 95% are in the brackets.

Effort-change (R1 to R2)	*Baseline*	*FutureInfo*	*PastInfo*	*PastWinUn*
OLS regression				
Confidence-change (R1 to R2)	0.09	0.02	0.06	0.04
	[0.01, 0.17]	[−0.06, 0.1]	[0, 0.12]	[−0.04, 0.11]
Winning probability in R1	−17.96	−−	−14.03	−5.14
	[−24.17, −11.75]		[−21.95, −6.12]	[−15.91, 5.63]
Constant	9.81	<0.01	6.99	1.10
	[6.39, 13.23]	[−1.27, 1.27]	[2.88, 11.1]	[−4.5, 6.69]
IV regression				
Confidence-change (R1 to R2)	0.20	0.08	0.15	0.03
	[0.05, 0.36]	[−0.13, 0.29]	[0, 0.3]	[−0.27, 0.32]
Winning probability in R1	−18.75	–	−13.87	−4.90
	[−25.08, −12.42]		[−21.88, −5.86]	[−17.43, 7.63]
Constant	10.55	−0.15	7.01	0.95
	[7, 14.1]	[−1.5, 1.2]	[2.85, 11.17]	[−5.89, 7.79]

which "hinders adept execution of acquired capabilities" (Bandura (1982)). On the other hand, the confidence gained from success can lead to an "altered and felt state of mind in which a performer senses things going unstoppably his or her way." This enables him/her "to perform at a level not ordinarily possible" (Iso-Ahola and Dotson (2014)). In economics, Compte and Postlewaite (2004) have discussed the implications of such a psychological effect to understand performance.

One possibility is that the mechanisms underlying such an effect are purely psychological/physiological in a way which do not lay themselves to an analysis in terms of rational strategy. Another possibility could be, however, that this "psychological momentum" is the result of a rational strategy in contests, conceived as games of incomplete information, where players do not know their relative strength initially. As the game progresses, they get information about their relative strength. In particular, after an initial win, players learned they are more likely to be the strongest of the two players. This information could intuitively lead winners to expend more effort as future success is perceived more likely.

This possible mechanism calls for the investigation of dynamic contests as games of incomplete information. Until now, only a few specific cases of dynamic contests have been solved in a setting with incomplete information (Konrad (2009), Münster (2009), Ederer (2010), Kubitz (2015), Miklós-Thal and Ullrich (2016)). The unknown heterogeneity of players adds a layer of complexity to the game: it not only has to be solved by backward induction starting from the end, but the equilibrium strategy at the end of the contest depends on the Bayesian beliefs formed from the *equilibrium strategies* and outcomes in all the previous rounds. Given that earlier outcomes influence later beliefs and later strategies, some nontrivial strategies can emerge in such settings, such as sandbagging whereby contestants who believe they are strong expend less effort to bias the beliefs of the other player (Münster (2009), Kubitz (2015)).

We should note that, however, some aspects of our results raise some doubts about the momentum we observe being the result of a fully rational strategy. First, we found that there is no strategic momentum, even though we would expect it to exist if players follow rational strategies. Second, we find that most participants failed at doing the hardest backward induction game in the post-experiment questionnaire (only 11% found the solution). This game only features simple binary comparisons of outcomes at four nodes of a simple game tree. It is allegedly easier to solve than a best-of-three contest, which requires to identify the equilibrium strategy at each round to identify its (expected) payoffs. Third, we find that players update positively their beliefs about their chances of winning in the *FutureInfo* treatment even though winning (randomly) in round 1 has no significant effect on their chances of winning in round 2. It raises doubts on the ability of participants to form accurate Bayesian beliefs. Nonetheless, rational strategies in dynamic contests with incomplete information may be worth investigating to improve our understanding of actual behavior in contests.

7. Conclusion

We find a clear positive causal effect of winning in a dynamic contest. This effect is not generated by the standard game-theoretic explanation, which identifies that an initial outcome creates an asymmetry in incentives between the leader and the laggard. Our results point instead to the effect of winning on the players' confidence. We observe that winning increases players' relative performances when it increases their beliefs in their relative strength. On the contrary, when players are fully informed about their performance—and, therefore, winning does not bring them any additional information—we do not observe a positive effect of an early success.

The main result of our study is that, whether for fully behavioral or for strategic reasons, players' belief updating seems to be the driving factor behind the momentum effect we observe, unlike what is predicted by most game-theoretic models of dynamic contests.

Understanding how competitors form and update their beliefs about their relative strength is critical to understand their motivation and strategies. Whether these beliefs are accurate or not, they may be influenced by past success and the updated beliefs may influence future performance. Even with an initially even playing field, subjective self-confidence can play a critical role in future performance and, therefore, contribute to putting identical people on different paths in terms of long term success.

References

Abadie, A. and G. W. Imbens (2008), "On the failure of the bootstrap for matching estimators." *Econometrica*, 76 (6), 1537–1557. [369]

Apesteguia, J. and I. Palacios-Huerta (2010), "Psychological pressure in competitive environments: Evidence from a randomized natural experiment." *American Economic Review*, 100 (5), 2548–2564. [356]

Araujo, F. A., E. Carbone, L. Conell-Price, M. W. Dunietz, A. Jaroszewicz, R. Landsman, D. Lamé, L. Vesterlund, S. W. Wang, and A. J. Wilson (2016), "The slider task: An example of restricted inference on incentive effects." *Journal of the Economic Science Association*, 2 (1), 1–12. [362]

Bandura, A. (1982), "Self-efficacy mechanism in human agency." *American psychologist*, 37 (2), 122. [357, 380]

Berger, J. and D. Pope (2011), "Can losing lead to winning?" *Management Science*, 57 (5), 817–827. [356]

Bergerhoff, J. and A. Vosen (2015), "Can being behind get you ahead? Reference dependence and asymmetric equilibria in an unfair tournament." Bonn Econ Discussion Papers. [356]

Binmore, K., J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked (2002), "A backward induction experiment." *Journal of Economic theory*, 104 (1), 48–88. [375]

Chen, D. L., M. Schonger, and C. Wickens (2016), "oTree—an open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance*, 9, 88–97. [367]

Cohen-Zada, D., A. Krumer, and Z. Shtudiner (2017), "Psychological momentum and gender." *Journal of Economic Behavior & Organization*, 135, 66–81. [357]

Compte, O. and A. Postlewaite (2004), "Confidence-enhanced performance." *The American Economic Review*, 94 (5), 1536–1557. [380]

Dechenaux, E., D. Kovenock, and R. Sheremeta (2015), "A survey of experimental research on contests, all-pay auctions and tournaments." *Experimental Economics*, 18 (4), 609–669. [362]

Descamps, A., C. Ke, and L. Page (2022), "Supplement to 'How success breeds success'." *Quantitative Economics Supplemental Material*, 13, https://doi.org/10.3982/QE1679. [369]

Ederer, F. (2010), "Feedback and motivation in dynamic tournaments." *Journal of Economics & Management Strategy*, 19 (3), 733–769. [361, 380]

Erkal, N., L. Gangadharan, and B. Han Koh (2018), "Monetary and non-monetary incentives in real-effort tournaments." *European Economic Review*, 101, 528–545. [362]

Fan, J. (1992), "Design-adaptive nonparametric regression." *Journal of the American statistical Association*, 87 (420), 998–1004. [369]

Ferrall, C. and A. A. Jr. Smith (1999), "A sequential game model of sports championship series: Theory and estimation." *The Review of Economics and Statistics*, 81 (4), 704–719. [356]

Fu, Q., C. Ke, and F. Tan (2015), "“success breeds success” or “pride goes before a fall”?: Teams and individuals in multi-contest tournaments." *Games and Economic Behavior*, 94, 57–79. [356, 366]

Fu, Q., J. Lu, and Y. Pan (2015), "Team contests with multiple pairwise battles." *The American Economic Review*, 105 (7), 2120–2140. [357, 359]

Gächter, S., L. Huang, and M. Sefton (2016), "Combining "real effort" with induced effort costs: The ball-catching task." *Experimental economics*, 19 (4), 687–712. [362]

Gauriot, R. and L. Page (2019), "Does success breed success? A quasi-experiment on strategic momentum in dynamic contests." *Economic Journal*, 129 (624), 3107–3136. [355, 356, 361]

Gill, D. and V. Prowse (2012), "A structural analysis of disappointment aversion in a real effort competition." *American Economic Review*, 102 (1), 469–503. [355]

Gill, D. and V. Prowse (2014), "Gender differences and dynamics in competition: The role of luck." *Quantitative Economics*, 5 (2), 351–376. [368]

Grabiszewski, K. and A. R. Horenstein (2019), "Skills, complexity, and backward induction." Available at SSRN 3337169. [367, 376]

Greiner, B. (2015), "Subject pool recruitment procedures: Organizing experiments with ORSEE." *Journal of the Economic Science Association*, 1 (1), 114–125. [367]

Harris, C. and J. Vickers (1987), "Racing with uncertainty." *The Review of Economic Studies*, 54 (1), 1–21. [357]

Heckman, J. J., H. Ichimura, and P. Todd (1998), "Matching as an econometric evaluation estimator." *The Review of Economic Studies*, 65 (2), 261–294. [369]

Hollard, G., S. Massoni, and J.-C. Vergnaud (2016), "In search of good probability assessors: An experimental comparison of elicitation rules for confidence judgments." *Theory and Decision*, 80 (3), 363–387. [367]

Huck, S., N. Szech, and L. M. Wenner (2015), "More effort with less pay: On information avoidance, belief design and performance." CESifo Working Paper Series No. 5542. [361]

Iso-Ahola, S. E. and C. O. Dotson (2014), "Psychological momentum: Why success breeds success." *Review of General Psychology*, 18 (1), 19–33. [360, 380]

Iso-Ahola, S. E. and K. E. N. Mobily (1980), "“psychological momentum”: A phenomenon and an empirical (unobtrusive) validation of its influence in a competitive sport tournament." *Psychological Reports*, 46 (2), 391–401. [357]

Johnson, E. J., C. Camerer, S. Sen, and T. Rymon (2002), "Detecting failures of backward induction: Monitoring information search in sequential bargaining." *Journal of Economic Theory*, 104 (1), 16–47. [375]

Klumpp, T. and M. K. Polborn (2006), "Primaries and the New Hampshire effect." *Journal of Public Economics*, 90 (6), 1073–1114. [357]

Kocher, M. G., M. V. Lenz, and M. Sutter (2012), "Psychological pressure in competitive environments: New evidence from randomized natural experiments." *Management Science*, 58 (8), 1585–1591. [356]

Konrad, K. A. (2009), *Strategy and Dynamics in Contests*. Oxford University Press. [360, 361, 380]

Konrad, K. A. and D. Kovenock (2006), "Equilibrium and efficiency in the tug-of-war." SFB/TR 15 Discussion Paper. [357]

Konrad, K. A. and D. Kovenock (2009), "Multi-battle contests." *Games and Economic Behavior*, 66 (1), 256–274. [357, 359, 361]

Kubitz, G. (2015), *Repeated Contests With Private Information*. New York University, New York. [361, 380, 381]

Mago, D. Shakun, and L. Razzolini (2019), "Best-of-five contest: An experiment on gender differences." *Journal of Economic Behavior & Organization*, 162, 164–187. [361]

Mago, S. D., R. M. Sheremeta, and A. Yates (2013), "Best-of-three contest experiments: Strategic versus psychological momentum." *International Journal of Industrial Organization*, 31 (3), 287–296. [355, 361]

Mago, S. D. and R. M. Sheremeta (2019), "New Hampshire effect: Behavior in sequential and simultaneous multi-battle contests." *Experimental Economics*, 22, 325–349. [355]

Malueg, D. A. and A. J. Yates (2010), "Testing contest theory: Evidence from best-of-three tennis matches." *The Review of Economics and Statistics*, 92 (3), 689–692. [355, 361]

Markman, K. D. and C. L. Guenther (2007), "Psychological momentum: Intuitive physics and naive beliefs." *Personality and Social Psychology Bulletin*, 33, 800–812. [357]

Merton, R. K. (1968), "The Matthew effect in science." *Science*, 159 (3810), 56–63. [355]

Miklós-Thal, J. and H. Ullrich (2016), "Career prospects and effort incentives: Evidence from professional soccer." *Management Science*, 62 (6), 1645–1667. [361, 380]

Miller, J. Benjamin, and A. Sanjurjo (2018), "Surprised by the gambler's and hot hand fallacies? A truth in the law of small numbers." *Econometrica*, 86 (6), 2019–2047. [355, 357]

Münster, J. (2009), "Repeated contests with asymmetric information." *Journal of Public Economic Theory*, 11 (1), 89–118. [361, 380, 381]

Nickell, S. (1981), "Biases in dynamic models with fixed effects." *Econometrica*, 1417–1426. [368]

Paserman and M. Daniele (2010), "Gender differences in performance in competitive environments? Evidence from professional tennis players." Working paper. [356]

Robertson, I. H. (2012), "The winner effect: The neuroscience of success and failure." Macmillan. [355]

Smith, J. A. and P. E. Todd (2005), "Does matching overcome LaLonde's critique of non-experimental estimators?" *Journal of Econometrics*, 125 (1), 305–353. [369]

Todd, P. E. (2010), "Matching estimators." In *Microeconometrics* (S. N. Durlauf and L. E. Blume, eds.), 108–121, London: Palgrave Macmillan UK. [369]

Tong, K. and K. Leung (2002), "Tournament as a motivational strategy: Extension to dynamic situations with uncertain duration." *Journal of Economic Psychology*, 23 (3), 399–420. [356]

Tullock, G. (1980), "Efficient rent seeking." In *Towards a Theory of the Rent-Seeking Society* (J. M. Buchanan, R. D. Tollison, and G. Tullock, eds.), 97–112, Texas AM University Press. [356]

Van de Rijt, A., S.M. Kang, M. Restivo, and A. Patil (2014), "Field experiments of success-breeds-success dynamics." *Proceedings of the National Academy of Sciences*, 111 (19), 6934–6939. [356]

Co-editor Peter Arcidiacono handled this manuscript.

Manuscript received 8 July, 2020; final version accepted 13 July, 2021; available online 24 August, 2021.