Journal of Economic Psychology 104 (2024) 102753

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

The leadership fallacy: How misattribution of leadership leads to a blaming game[☆]

Nikola Frollová[a,*], Marcel Tkáčik[a], Petr Houdek[a,b]

[a] Prague University of Economics and Business, Faculty of Business Administration, Winstona Churchilla 1938/4, 120 00 Praha 3-Žižkov, Czechia
[b] Charles University, Faculty of Social Sciences, Smetanovo nábřeží 6, 110 01 Praha 1, Czechia

A R T I C L E I N F O

Keywords:
Attribution
Overconfidence
Leader-centered
Attribution biases
Self-serving bias

A B S T R A C T

Assigning responsibility for a project's success or failure is key to organizational performance, yet attribution fallacies often interfere. Our experimental study (*N*=339) shows team members mistakenly attribute too much influence to their leaders on task outcomes. Despite task outcomes being randomly determined by easy or hard difficulty rather than leadership, leaders received undue credit or blame. Leaders assessed their teams more negatively in difficult tasks, except for female leaders, who were more lenient in assessing both conditions than men. Leaders' self-assessments did not differ between experimental conditions, confirming their self-motivated evaluation; moreover completing an easy task boosted their confidence for harder challenges. Our study shows that attributional errors manifest differently in the evaluation of leaders and followers and demonstrates that success in simpler tasks can increase leaders' confidence, potentially leading to riskier behaviors.

1. Introduction

An organization's performance is often explained through its top executives' experience, skills, and personalities (Hambrick, 2007; Loerakker & van Winden, 2017). An individual's success is then perceived as a consequence of effort, skill, and personal traits such as talent or intelligence (Lovell & Kosten, 2000; Pluchino et al., 2018; Ridgell & Lounsbury, 2004). As the Romance-of-Leadership approach highlights, the management or leadership of an organization is generally viewed as the only or main influence on its performance (Bligh et al., 2011; Gentry & Sparks, 2012; Meindl & Ehrlich, 1987); and followers may remain loyal to a leader's guidance even when based on inconsequential information, sometimes to their own detriment (Cavalcanti et al., 2023). Moreover, random and situational factors affecting performance are often overlooked, even if they play a significant role in achieving success (Liu & de Rond, 2016).

People sometimes admit that a certain degree of luck can affect their success. However, the role of luck is underestimated. In many situations, leaders' behaviors and decisions make essentially no difference to organizational outcomes (Bligh et al., 2011; Porter & McLaughlin, 2006; Wageman & Richard Hackman, 2007). A stream of literature highlights random and situational forces as a more realistic explanation for differences in performance between individuals and organizations (Harrison & March 1984; Houdek, 2017; Liu & de Rond, 2016).

☆ Data Availability Statement: In adherence to open science principles, datasets and experimental protocols supporting this article's conclusions are accessible via the Open Science Framework (OSF) repository at the following link: https://osf.io/z2nwr/.

* Corresponding author.
E-mail address: nikola.frollova@gmail.com (N. Frollová).

https://doi.org/10.1016/j.joep.2024.102753
Received 11 October 2022; Received in revised form 2 May 2024; Accepted 29 July 2024
Available online 5 August 2024

Are leaders and their followers aware of the impact of chance on their performance? Are they able or willing to reflect this influence in their decision-making when allocating responsibility for a project's success or failure? Seeing how chance plays an important role in individual and organizational outcomes can be psychologically challenging because it diminishes the feelings of agency and control (Langer, 1975; Meissner & Wulf, 2017). Several cognitive biases provide an illusion of control that allows people to perceive the impact of their behavior on the outcomes even when this causality is missing (Lassiter et al., 2002). In negative circumstances with less perceived control, the tendency to perceive illusory patterns is even stronger (Quan et al., 2011), arguably as a consequence of the desire to see the world in a predictable state (Whitson & Galinsky, 2008). Thus, on the one hand, biases may be useful coping mechanisms when faced with uncertainty; on the other hand, biases prevent people from seeing the underlying mechanisms behind an outcome, which can lead to poor decisions in the long term (Williams, 2021).

For every organization, there is an imperative to understand the determinants of success and failure. Stakeholders are motivated to determine why tasks and processes fail and try to set preventive measures effectively and targeted. However, there is an asymmetry in attribution between success and failure (Liu & de Rond, 2016). The source of failure can be easily misattributed due to discounting one's role in a failure (Jones & Davis, 1965; Nasir Zadeh et al., 2020), followed by defensiveness bias and neglecting liability in the forms of excuses; on the other hand, people tend to place themselves as the key agents of positive events (Standing et al., 2006). None of these polarized attributions serves as a lesson for improving the organization's processes because the causes of failure or success are not correctly identified.

Simon's procedural rationality (Khalil, 2022) acknowledges that individuals cannot optimize decision-making due to cognitive constraints imposed by the complexity of the world. As such, individuals rely on satisficing strategies settling for an acceptable solution so that the cognitive costs of the decision are not prohibitive. This perspective is critical in understanding how organizational actors may interpret outcomes and attribute them to the actions of colleagues or leaders, often ignoring difficult to detect counterfactuals (Jehiel, 2018). This otherwise effective procedure can be a significant factor contributing to organizational failures (Garicano & Rayo, 2016), especially in fast-changing situations where real-time coordination between leaders and followers is needed (Milgrom & Roberts, 1988) and interaction dynamics between them can be a source of biases (Bligh et al., 2011).

Our study addresses the call to determine how random situational factors impact the evaluation of leaders and followers (Bligh et al., 2011). Secondly, we enrich the existing literature by adding multiple rounds of an experimental game parallel to organizational learning to examine the consequences of incorrect attribution of merit. Finally, we respond to the demand for leadership experiments (Avolio et al., 2009; Kirkpatrick & Locke, 1996), and for experiments in organizational theory in general (Schilke et al., 2019).

2. Research questions

1. How much is the assessment of the leaders dependent on random external factors beyond their control?
2. Do the same rules apply to self-assessments by leaders, assessments of leaders by followers, and assessments of followers by leaders?
3. Will a random success or failure affect the leader's further actions?

The first part of this pre-registered study is a conceptual replication of team members' tendencies to overestimate a leader's influence on their performance based on the group size (Weber et al., 2001). Using conceptual replication, we want to verify the original findings through several methodological improvements and higher statistical power (Köhler & Cortina, 2019). The results will allow us to assess the contextual sensitivity of the original findings and help us understand how the attribution of leadership ability can depend on the task's difficulty, where it is possible to learn and adapt. (The original study's design does not allow for participants' learning.) In addition, our design allows us to see how an external factor affects leaders' self-evaluation; we are also interested in how they evaluate team members.

The second part of the experiment focuses on leaders' decision-making. We were interested in whether a difference in a situational factor can affect the willingness of leaders to solve a task under new conditions, with different difficulties and rewards. When previous success was determined by the external factor, i.e., under the easy-task conditions in the first part of the experiment, we predict that these leaders will be more willing to pay for more challenging, higher-rewarded tasks compared to the leaders who previously completed a task under more challenging conditions.

3. Theory and hypotheses

3.1. Attribution error, leaders and followers

The factors that influence the outcome of an individual's or team's performance could be explained by internal factors, such as skill and effort, or external factors, such as the difficulty of the task, situational effects, and sometimes (good or bad) luck. Nevertheless, individuals often struggle to account for unseen (counterfactual) factors in their decision-making, largely ignoring external influences that are not immediately apparent or directly observable to them (Enke, 2020). They often develop an incorrect mental representation and judge or behave based on misattribution, leading to undeserved blame and rewards (Liu & de Rond, 2016). Moreover, research has highlighted the connection between such mistakes in causal reasoning and the development of overconfident beliefs (Jehiel, 2018). Determining the importance of internal or external effects on an outcome is affected by overattribution of individual traits and behavior rather than situational factors, known as fundamental attribution error (Ross et al., 1977). In companies or teams, attribution error is manifested by identifying the leader as the main cause of collective performance (Swift et al., 2013; Wageman & Richard Hackman, 2007). There is evidence that managers are evaluated based on outcomes — either positive or negative — even though their decision-making process could be identical (Dillon & Tinsley, 2008). This is also present in attributing a company's performance to its top executives (Bertrand & Mullainathan, 2001; Meindl & Ehrlich, 1987) or basing attribution of leaders' characteristics just on salient information (Ensari & Murphy, 2003). Another attribution bias leading to the overestimation of leadership is the illusion of control (Langer, 1975; Meissner & Wulf, 2017), leading to blaming leaders for failures outside their control (Salancik & Meindl, 1984).

On the other hand, a specific subtype of attribution error related to the leadership position is self-serving bias. Leaders credit good results to themselves but low group performance to followers (Dobbins & Russell, 1986; Erkal et al., 2022). There is also evidence that self-attribution has a cumulative effect over time (Libby & Rennekamp, 2012) as leaders sacrifice innocent team members as scapegoats in failure cases. Heidhues et al. (2018) showed that this misattributing of success or failure can lead to unrealistic expectations and hinder effective group learning. In line with these findings, Coutts, Gerhards, and Murad (2020) demonstrated that individuals tend to attribute success to their own abilities while attributing failure to external factors or other group members, thus maintaining a self-serving bias that protects their self-esteem and preserves positive self-perceptions. Similarly, Hestermann and Le Yaouanq (2021) highlighted how individuals engage in self-serving attributions to maintain a positive self-image and enhance their social status within the group.

Differences in the misatribution of causes also depend on various factors; for example, leadership style (Lopez & Ensari, 2014), culture (Yazici, 2011), or gender (Andrews, 1987; Elsesser, 2016; Simon & Feather, 1973) – women tend to attribute their failures more to themselves and success to external factors, while men have the opposite tendencies.

The mechanism of incorrect attribution of success or failure to leaders was documented in a laboratory coordination study by Weber et al. (2001). Participants were randomly assigned to two conditions: pairs of participants (a small group) or a group of ten people (a big group). Both groups solved the same coordination game. The success of the teams depended on their ability to coordinate. A leader was randomly selected for each group. Leaders read a useless message to the group that aimed to improve its coordination. The pairs were more successful than the big groups because of much easier coordination. Due to the attribution error, team members rated the leader more positively in pairs than in large groups (although the condition affected outcomes).

In our study, we assume that under the effect of an attribution error, people in teams will tend to evaluate their leaders according to the difficulty of the conditions, just like in the original experiment (Weber et al., 2001) even though the leader's action, operationalized as useless advice, will be the same for all conditions. Since we are not looking at the effect of the leadership style but the effect of an external factor on the evaluation in comparison to the original experiment, we operationalized the task difficulty as a Numberlink game (Ahle, 2012/2020) with varying difficulty. This design modification aims at examining the contextual sensitivity of the original findings. Because it is an effort task with the possibility of learning, it also has greater external validity for organizational settings. Hypotheses 1a and 1b are conceptual replications of Weber et al.'s findings.

Hypothesis H1a. Participants in the easy condition will judge leaders more positively than participants in the difficult condition.
Hypothesis H1b. The evaluation of the team and leader will be associated with the condition – the evaluation of both the leader and the team will be higher in the easy condition.

3.2. Overconfidence based on misattribution

Misattribution can lead to a significantly skewed assessment of abilities and skills. People tend to overestimate the probability of failure and success compared to neutral events. This is because they incorrectly attribute the perceived arousal caused by the saliency of the outcome as evidence for a higher likelihood of a particular event in the future (Vosgerau, 2010). This misattribution leads to inappropriate risk assessment and over- or under-confidence in their own performance (Fabricius & Büttgen, 2015).

Overconfidence in performance estimation has been observed in a variety of domains; in managers' earnings forecasts (Libby & Rennekamp, 2012); stock market bubbles (Camerer & Lovallo, 1999); economic forecasts (Braun & Yaniv, 1992) or prediction of sports outcomes (Ronis & Yates, 1987). Although most executives expect that other executives overestimate how much they contribute to the company's performance, they show the same overconfidence themselves (Libby & Rennekamp, 2012). We assume that attribution error caused by ignoring external factors leads to over-learning from successes resulting in overconfidence (Camerer & Lovallo, 1999; Fabricius & Büttgen, 2015; Libby & Rennekamp, 2012). Indeed, overconfidence may provide a signal of competence and reduce social stress (Ronay et al., 2019). It is also connected with extraversion (Schaefer et al., 2004), sensation seeking (Grinblatt & Keloharju, 2009), or gender; men tend to be more overconfident than women (Barber & Odean, 2001).

We also examine whether participants will make different decisions from the rest of the team after a randomly assigned leadership role. The power-approach theory (Keltner et al., 2003) suggests that a position of authority (a leader) can increase power because it provides control over resources and the ability to influence others to achieve and accomplish something (Lunenburg, 2012). Possessing power may lead individuals to experience more positive and less negative affect and to enjoy higher self-esteem (Anderson et al., 2012). This mechanism can reinforce the effects of attribution error, leading to overconfidence.

We expect that situational success will lead to overconfidence in a leader, and it will be displayed as a leader's willingness to pay to participate in a task where there is a lower probability of achieving the initial success because that success was caused by a random situational factor: low task difficulty.

Hypothesis H2a. Leaders from the easy condition will be more willing to pay for more difficult versions of the task than leaders from the difficult condition.
Hypothesis H2b. Group willingness to pay (willingness to pay averaged over all group members) will be different between conditions; willingness to pay will be higher in groups playing the easy condition.

Hypothesis H2c. Willingness to pay will be associated with a position in the group – the leader of the group will exhibit a higher willingness to pay than followers in the same group.

We conducted an experimental study to test our hypotheses, divided into two phases. In phase 1, teams play cognitive game tasks of two arbitrarily different difficulties to examine hypotheses 1a and 1b and conceptually replicate the study of Weber et al. (2001). We predict that Team Evaluations, Leader's Evaluation by team members, and Leader's Self-Evaluation will be higher in the easy condition. To examine hypotheses 2a, 2b, and 2c, we observe the willingness of the leaders to pay for playing another version of the task in phase 2. After the experience of success or failure associated with the condition, we predict that Willingness to Pay for a chance to play an additional game in three options of difficulty will be influenced by that experience; more specifically, that leaders will be willing to pay higher amounts to enter into a difficult game under the influence of success given by the situational factors.

4. Methods

4.1. Participants

We recruited 339 participants; they were invited some time in advance, so the final sample size differed slightly from the planned 300. A power analysis determined the sample size; the goal was reaching 0.80 power with alpha level set to 0.05 and standardized effect size Cohen's d = 0.50. Participants were recruited from the Laboratory of Experimental Economics (LEE) participants pool, consisting mostly of Czech university students. The analysis was performed with data from the 339 participants, of whom 45% were women. The mean age of the participants was 24.2 years (M=24, SD=3.6). All participants completed the experiment, and we did not exclude anyone from the analysis.

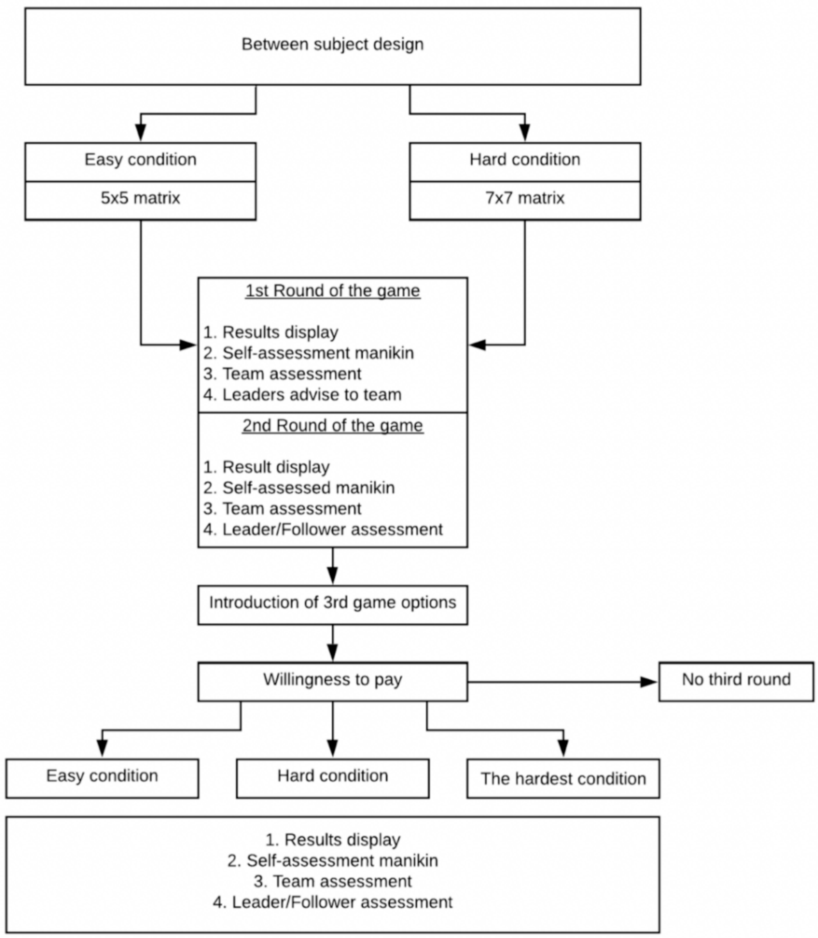

Fig. 1. Overview of experimental design.

4.2. Procedure

Participants were told that the study would take 45 minutes and that the reward would depend on their performance in the task. The show-up fee was 100 CZK (approx. 4.50 USD), the maximum reward was 400 CZK (approx. 17.50 USD), and the mean reward was 300 CZK (approx. 13 USD). The participants filled in a consent form and were randomly assigned to a three-person group, either to the role of a leader or a follower (team members). Teams were randomly assigned either to the easy or difficult task conditions in a game called Numberlink (Ahle, 2012/2020). Numberlink is a logic puzzle involving finding paths to connect numbers in a grid. The difficulty of the conditions is operationalized using the size of the game matrix (easy condition: 5 × 5 matrix; difficult condition: 7 × 7 matrix).

In the first round of the game (see Fig. 1 for an overview of the study), all participants, including the leaders, had the same instruction: five minutes to solve as many rounds of the game as possible, from ten rounds in total. For each resolved game in the round, there were the following rewards:

Easy condition (5 × 5 matrix): 8 CZK (0.40 USD) for every team member, i.e. 24 CZK (1.20 USD) for the team.

Hard condition (7 × 7) matrix: 16 CZK (0.80 USD) for every team member, i.e. 48 CZK (2.20 USD) for the team.

The sum of the rewards was displayed after each round. After the first round, all participants were informed of the results of the team's effort. Participants saw their results, but also the results of colleagues in their team. Team remuneration was based on the average performance of the team members. The team remuneration was distributed evenly among the players. After the first game round, participants reported their emotions on the Self-Assessment Manikin (SAM) scale to detect their feelings from failure or success. We assumed that people from an easy-condition task would report more positive emotions (Kassas et al., 2022). Then players evaluated team performance and their own contribution to the team's outcome.

Before the second round, the leaders were challenged to choose one of two prepared pieces of advice to be sent to the team. Each piece of advice gave the team members information about how to approach the game. There were actually three similar-sounding pieces of advice; however, only two were randomly selected in order to avoid the effect of the advice (all pieces of advice were perceived as equally informative at the pretest, and the order of pieces of advice was randomized). The second round was identical to the first round, along with an assessment of emotions and evaluation of team members. In addition, there was a question on the leader's contribution and self-evaluation.

Before the third round, leaders were prompted to decide how much they were willing to pay for their teams to take part in the round. The conditions of the third round introduced to all participants were the following: a 5 × 5 matrix or a 7 × 7 matrix (the same as in the first and second rounds), or an 8 × 8 matrix. The reward mechanism of already-known games was the same as in the previous rounds. The 8 × 8 difficulty was the highest and most rewarded. For each resolved game in the round, there was a reward of 24 CZK (1.20 USD) for every team member (72 CZK (3.20 USD) for the whole team).

The Becker–DeGroot–Marschak (BDM) method (James, 2007) was used for eliciting willingness to pay to participate in the third round. In the BDM method, participants formulate bids, which are compared to a number determined by a random number generator. When the participants' bids are greater than the generated number, they "win", and a third round takes place in which they pay the value of the generated number. If the participants' bids are lower than the generated number, they "lose", and the third round will not take place, in which case the participant pays nothing. The bid value thus indicates the preference to participate in the next round. Participants were introduced to the BDM mechanism. Even though participants knew the actual decision would be based solely on the leaders' willingness to pay, we asked everyone for this measure. The price (their winning bid) for taking part in the third round was to be evenly split among all three team members if the round was played. After participants stated their willingness to pay for each condition, one of the three conditions was randomly selected, and the BDM method determined whether or not each team would perform a task in the third round.

The mechanism and all conditions were explained, and the participants' understanding was checked before they stated their willingness to pay. A third round was followed by a final (self-)evaluation of the team and the leader, demographics, Self-Efficacy (Schwarzer & Jerusalem, 1995) and Sense of Power scales (Anderson et al., 2012) to see how leaders felt about the result achieved. A post-experimental questionnaire about the participants' decision-making process and payment of remuneration concluded the experiment.

5. Overview of experimental design

5.1. Manipulations

The experiment employed a between-subject design in a laboratory setting with two factors. The first one was game difficulty (hard and easy), and the second factor was the position in the group (leader or follower). Just like Weber et al. (2001) we operationalized leadership as a simple, useless address by one randomly selected group member. The operationalization allows us to follow the effect of external factors without the influence of any aspects of true leadership skills or effort.

5.2. Measures

The main dependent variables of interest are ratings of the leader by team members and the leader's self-rating. Other dependent variables were satisfaction with results, team evaluation, and willingness to pay. We also control for the following variables: gender, age, education, beliefs, life satisfaction (Cheung & Lucas, 2014), how important it is to be better than others, and how a participant thinks they stand in comparison to others financially (Kahneman et al., 2006). All measures and variables are summarized in Table 1..

5.3. Planned analysis

Regression models were used to estimate two main dependent types of variables (variable y_i): self-rating and the ratings of others by the leader and team; and willingness to pay for one of three games in the third round. We first estimate simple linear models with a variable for the condition and position in the group. We assume that every outcome y_i of player i is determined by the game condition and/or a game score. First, we regress dependent variables on the game condition alone and the game score alone. The models take these forms:

$$y_i = \beta_0 + \beta_1 * Hard + u \quad (1)$$

where β_0 is a constant, Hard is a dummy variable taking 1 if the group has been selected to play the hard condition and 0 otherwise, and u is a residual; and:

$$y_i = \beta_0 + \beta_1 * GameScoreR2 + u \quad (2)$$

where GameScore is a variable representing the number of matrices solved in the second round.

Due to the random assignment of participants to conditions, Equations (1) and (2) isolate and estimate the effects of the game condition and the game score from the second round on the outcome y_i, respectively, by using separate equations. To see into the combined influence of the game condition and the game score, we further specified a model that examines how the game score from the second round affects the dependent variable within its respective condition. This model is represented by:

$$y_i = \beta_0 + \beta_1 * Hard + \beta_2 * GameScoreR2 + \beta_3 * (Hard * GameScoreR2) + u \quad (3)$$

While our primary interest lies in differentiating between the effects of the game condition and individual scores, we further nuanced our approach for the third equation. Specifically, we estimated the third equation using a mixed-effects model to account for the inherent correlation between individual scores across the two rounds. This model permits random intercepts for each participant, acknowledging that an individual's scores from both rounds are not independent events. By incorporating these random intercepts, we can account for the individual-level variation that might arise. Meanwhile, fixed effects in the model elucidate the influence of the game condition and its interaction with scores from the second round.

Our methods and predictions were pre-registered on the Open Science Framework (https://osf.io/z2nwr/). The experiment was approved by the Ethics Committee of the Prague University of Economics and Business (Procol number 8/19).

Table 1
Overview of variables.

Variables	Rounds	Scale		Description	Source
Game score	1,2,3	0 (lowest) to 10 (highest)		Score from the round	Own
Satisfaction of results	1,2,3	1 to 9	(very dissatisfied) (very satisfied)	Satisfaction with the last round's results	Own
Player self-evaluation	1,2,3	1 to 9	(very dissatisfied) (very satisfied)	Satisfaction with own. contribution to the team in the last round	Own
Team evaluation	1,2,3	1 to 9	(very dissatisfied) (very satisfied)	Assessment of satisfaction and contribution of team players	Weber et al. (2001)
Leader self-evaluation	2,3	1 to 9	(poorly) (perfectly)	Leader self-evaluation	Weber et al. (2001)
Willingness-to-pay	3	0 to 200 CZK		Becker-Degroot-Marschak mechanism	James (2007)
Self-assessment manikins – feelings	1,2,3	0 to 9	(happy) (sad)	Pleasure, measuring a person's affective reaction to the game and its results.	Bradley and Lang (1995)
Self-assessment manikins – intensity of emotions	1,2,3	0 to 9	(highest intensity) (lowest intensity)	Arousal measures a person's affective reaction to the game and its results.	Bradley and Lang (1995)
Self-assessment manikins – dominance	1,2,3	0 to 9	(least powerful) (most powerful)	Dominance measures a person's affective reaction to the game and its results.	Bradley and Lang (1995)
Sense of power	Questionnaire	1 to 6	(not at all true) (exactly true)	Explanation in the on–line appendix	Anderson et al. (2012)
Self-efficacy	Questionnaire	1 to 7	(disagree strongly) (agree strongly)	Explanation in the on–line appendix	Schwarzer & Jerusalem (1995)

5.4. Pretest

We pretested that there was no problem in achieving a total success rate of solving all the matrices (10 out of 10) in the easy-task condition. On the other hand, the hard condition of the game is more challenging.

6. Results

In the first round, teams assigned to the hard condition solved on average (M=2.85, SD=0.97) 64% fewer matrices than teams playing in the easy condition (M=8.04, SD=2.68), $t(337) = 22.49$, $p < 0.001$. These results also serve as a check for the experimental manipulation. Because the reward per matrix in the hard condition paid only twice as much (24 CZK for every easy game vs 48 CZK for a hard game), the winnings of players in the hard condition (M=46.5, SD=8.59) are in the first round 28% lower compared to the those of the players playing in the easy condition (M=64.63, SD=11.13), $t(337) = 16.41$, $p < .001$.

Players in the hard condition were 41% less satisfied with the first round results (M=3.39, SD=1.57) compared to the easy condition players (M=5.83, SD=2.54), $t(337) = 10.3$, $p < 0.001$. Similarly, hard condition players were 35.6% less satisfied with their contribution to the team (M=4.41, SD=2.77) than the easy condition players (M=6.85, SD=4.41), $t(337) = 8.76$, $p < 0.001$. Hard condition players were also 24% less satisfied with the work of other team members (M=4.87, SD=1.98) than easy condition players (M=6.46, SD=2.4), $t(337) = 6.56$, $p < 0.001$. Overall, the experimental manipulation successfully led to the differences in levels of satisfaction with personal and team performance (for the summary statistics and correlations, see Tables 2 and 3).

In the second round, after leaders' messages, the teams in both conditions experienced an improvement in their results: easy-condition players scored on average 10% better in the second round (M=8.85, SD=2.45) compared to the first round, $t(337) = -3.16$, $p < 0.001$. Hard condition players improved 31% in the second round (M=3.83, SD=2.83), $t(337) = -3.94$, $p < 0.001$ compared to the first round. The difference in improvement between the conditions is significant, $t(337) = -5.38$, $p < 0.001$. Nevertheless, the winnings of the players in the hard condition (M=62.1, SD=2.34) were, on average, still 12.7% lower compared to the players in the easy condition (M=71, SD=10.4), $t(337) = 4$, $p < 0.001$. Overall, in the first two rounds, players in the hard condition (M=108.5, SD=18.7) earned 20 % less compared to the players in the easy condition (M=135.6, SD=18.7), $t(337) = 9.8$, $p < 0.001$.

After the second round, participants answered questions about their satisfaction with the second round result, their contribution to the team, and the work of other team members. Hard condition players were again less satisfied with the second-round results (M=4.37, SD=2.08) compared to the easy-condition players (M=7.15, SD=2.59), $t(337) = 10.67$, $p < 0.001$. Similarly to the first round, players in the hard condition were 35.8% less satisfied with their contribution to the team (M=4.99, SD=2.59) than the players in the easy condition (M=7.78, SD=2.43), $t(337) = 10.2$, $p < 0.001$. Players in the hard condition were also 34.2% less satisfied with the work of other team members (M=5.07, SD=2.38) than the hard condition players (M=7.71, SD=2), $t(337) = 11.06$, $p < 0.001$. All these results confirm H1b. Compared to the first round, hard condition players were 29% more satisfied with their results, $t(296) = -4.57$, $p < 0.001$, 13% more satisfied with their own contribution to the team, $t(296) = -2.1$, $p < 0.001$, and insignificantly 4% more satisfied with the work of other team members, $t(296) = -0.82$, $p = 0.210$. Players in the easy condition were 23% more satisfied with their results, $t(378) = -5$, $p < 0.001$, 14% more satisfied with their own contribution to the team, $t(378) = -3.51$, $p < 0.001$, and 20% more satisfied with the work of other team members, $t(378) = -5.48$, $p < 0.001$.

Participants were then asked how well they thought the leader prepared their team for the last round. Followers playing the hard condition rated their leaders after the second round 27% worse (M=4.74, SD=2.08) than the team members in the easy condition (M=6.52, SD=2.19), $t(224) = 6.24$, $p < 0.001$) confirming H1a.

There was no difference in the self-assessment of leaders for both conditions: leaders in easy (M=5.67, SD=5.35) and hard conditions (M=5.35, SD=2.04) thought that they prepared their followers equally well for the second round, $t(111) = 0.82$, $p = 0.207$. However, while the leaders' self-assessment did not vary between the two conditions, leaders in the hard condition were 31% less satisfied with the contribution of team members (M=5.2, SD=2.54) than the leaders in the easy condition (M=7.59, SD=2.08), $t(111) = 5.5$, $p < 0.001$.

Table 4 shows how the leaders' evaluation by their followers and the leader's self-evaluation can be explained by the game difficulty and the game score. Models 3 and 7 are estimated using mixed-effects models with level equations specified as a score in the second round being influenced by the score in the first round and the score of the first round being influenced by the selection of the game condition (using random-effects). This helps to distinguish between the variance attributable to the selection of the game condition itself and the variance attributable to the team-specific performance within the condition. Model 2 shows that the rating of the leader by followers in both conditions increases by 0.38 points for every matrix solved (SE=0.05, $p < 0.001$). Accounting for the condition effect and the team-specific performance effect in Model 3, using the mixed-effect model to correct for the team performance inside the condition, the effect of the condition itself on the evaluation of the leader is 0.36 points for every matrix solved (SE=0.05, $p < 0.001$). Model 4 shows that followers increased the leader's evaluation equally in both conditions, $F(1.113) = 0.00$, $p = 0.968$), while model 5 shows the leaders in the hard condition did not evaluate themselves on average differently than in the easy condition, explaining the self-evaluation by the group score in the mixed-effects model 7, which shows that leaders increased their self-evaluation by 0.13 for every additional matrix solved ($b = 0.13$, SE=0.43, $p = 0.040$). OLS models are estimated with team-clustered robust standard errors and the mixed-effects models are estimated with robust standard errors.

6.1. How is the willingness to pay influenced by the game difficulty?

In the third round, all players could bid money to play either a 5x5 game, a 7x7 game, or an 8x8 game (with a maximum reward of

Table 2
Correlation Matrix.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Reward R1	1															
(2) Reward R2	0.27***	1														
(3) Reward R3	0.32*	0.79***	1													
(4) Score R1	0.75***	0.18***	0.26	1												
(5) Score R2	0.14***	0.59***	0.59*	0.35***	1											
(6) Score R3	0.32*	0.41***	0.47***	0.33***	0.31***	1										
(7) Leader self-evaluation, R2	0.01	0.31**	0.26	−0.07	−0.28	0.10	1									
(8) Satisfaction of results, R1	0.36***	−0.36***	−0.22	0.30***	−0.26***	−0.04	−0.11	1								
(9) Satisfaction w/ contribution to team, R1	0.38***	−0.15***	0.01	0.52***	−0.17***	0.22	0.15	0.59***	1							
(10) Satisfaction w/ work of other members, R1	0.06***	0.03***	0.09	−0.35***	−0.25**	−0.01	0.03*	0.23***	0.16***	1						
(11) Satisfaction of results, R2	0.17***	0.58***	0.39*	0.29***	0.63***	0.34***	−0.07	0.01***	−0.12***	−0.28***	1					
(12) Satisfaction w/ contribution to team, R2	−0.17***	0.20***	0.23	0.11***	0.71***	0.14***	−0.28	−0.05***	−0.14***	−0.34***	0.52***	1				
(13) Satisfaction w/ work of other members, R2	0.18***	0.42***	0.17	−0.17***	−0.27***	0.12	0.40***	−0.02***	−0.07***	0.45***	0.11***	−0.45***	1			
(14) Satisfaction of results, R3	0.30	0.33	0.45*	0.20	0.01	0.67***	0.42	0.08	0.26*	0.25*	0.12	−0.12*	0.38	1		
(15) Satisfaction w/ contribution to team, R3	0.15	0.07	0.18	0.23	0.13*	0.72***	0.01	0.11	0.27***	0.04	0.17*	0.22***	−0.10	0.69***	1	
(16) Satisfaction w/ work of other members, R3	0.28*	0.47*	0.43	−0.11	−0.04	0.17***	0.48*	0.06*	0.08	0.36***	0.15	−0.12	0.65***	0.28***	−0.10***	1

R corresponds to the game round (1,2,3). Correlation values computed by Spearman rank-order correlation coefficients.
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

Table 3
Summary statistics for the first and second round.

First round (R1).

		Overall			Easy condition				Hard condition			
Variables	N	Mean	Median	SD	N	Mean	Median	SD	N	Mean	Median	SD
Game Score R1	339	5.76	4	3.33	190	8.04	10	2.68	149	2.85	3	0.97
Team Reward R1	339	169.98	168	40.56	190	193.9	192	33.39	149	139.5	144	25.77
Satisfaction of results R1	339	4.76	4	2.48	190	5.83	6	2.54	149	3.39	3	1.57
Player self-evaluation R1	339	5.78	6	2.81	190	6.85	9	2.77	149	4.41	5	2.22
Team evaluation R1	339	5.76	6	2.36	190	6.46	7	2.4	149	4.87	5	1.98

Second round (R2)

		Overall			Easy condition				Hard condition			
Variables	N	Mean	Median	SD	N	Mean	Median	SD	N	Mean	Median	SD
Game Score R2	339	6.64	8	3.62	190	8.85	10	2.45	149	3.83	3	2.83
Team Reward R2	339	201.27	192	62.76	190	213.1	240	31.23	149	186.21	192	85.71
Satisfaction of results R2	339	5.93	6	2.75	190	7.15	9	2.59	149	4.37	4	2.08
Player self-evaluation R2	339	6.56	8	2.86	190	7.78	9	2.43	149	4.99	5	2.59
Team evaluation R2	339	6.55	7	2.53	190	7.71	9	2	149	5.07	5	2.38
Leader evaluation R2	226	5.72	5	2.31	126	6.5	7	2.19	100	4.74	5	2.08
Leader self-evaluation R2	113	5.53	5	2.08	64	5.67	5	2.12	49	5.35	5	2.04
Team reward R1+R2	339	371.26	384	85.8	190	407	408	56.02	149	325.7	336	95.3

Table 4
Evaluation after Round 2.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Variables	OLS	OLS	ME	OLS	OLS	OLS	ME	OLS
Hard	−1.76***				−0.32			
	(0.32)				(0.39)			
Group Score R2		0.38***	0.36**			0.13	0.13*	
		(0.05)	(0.05)			(0.07)	(0.43)	
Easy X Group Score R2				0.38**				0.2*
				(0.06)				(0.08)
Hard X Group Score R2				0.37**				0.38*
				(0.13)				(0.15)
Constant	6.5***	3.22***	3.32***	3.23***	5.67***	4.65***	4.65***	3.89***
	(0.21)	(0.36)	(0.33)	(0.53)	(0.27)	(0.5)	(0.43)	(0.69)
Observations	226	226	226	226	113	113	113	113
R-squared	0.14	0.23		0.23	113	113		113
Log.likelihood for ME			−479.53				−240.79	

OLS models estimated with team-clustered robust standard errors. Mixed-effects models estimated with robust standard errors. OLS columns correspond to OLS estimates, ME to mixed-effects estimates.
***p < 0.001, **p < 0.01, *p < 0.05.

720 CZK). Only bids by the leaders were taken as input to the bidding mechanism (see Table 5). Fig. 2 shows how the game difficulty influences the distribution of bids. As can be seen from the figure and is confirmed by the Shapiro-Wilk test for the normality, the distribution of bids does not follow the normal distribution; therefore we use Wilcoxon rank-sum tests for the analysis of bids. In accordance with H2a, leaders in the hard condition compared to the leaders in the easy condition bid 24.7% less on the 5x5 game (Mann-Whitney test, U=2.98, $p = 0.003$), 26.3% less on the 7x7 game (Mann-Whitney test, U=4.02, $p < 0.001$) and 23.4 % less on the 8x8 game (Mann-Whitney test, U=2.98, $p = 0.003$). Out of 114 groups, 68 groups were selected in a Becker-DeGroot-Marschak mechanism to play one of the games in the third round (60%). Of the groups, 41% came from difficult conditions and 59 % from

Table 5
Leaders' willingness to pay for the third round.

		Overall			Easy condition				Hard condition			
Variables	N	Mean	Median	SD	N	Mean	Median	SD	N	Mean	Median	SD
WTP for game 5x5	113	110.03	100	56.39	64	123.2	120	57.21	49	92.8	90	50
(% of team income from R1,R2)	113	29.64%	26%		64	30%	29%		49	28%	27%	
WTP for game 7x7	113	121.4	115	50.3	64	137	150	49.8	49	101.02	100	43.4
(% of team income from R1,R2)	113	32.7%	30%		64	34%	37%		49	31%	30%	
WTP for game 8x8	113	126.8	122	61.76	64	141.1	175	62.84	49	108.12	100	55.56
(% of team income from R1,R2)	113	34%	32%		64	35%	43%		49	33%	30%	

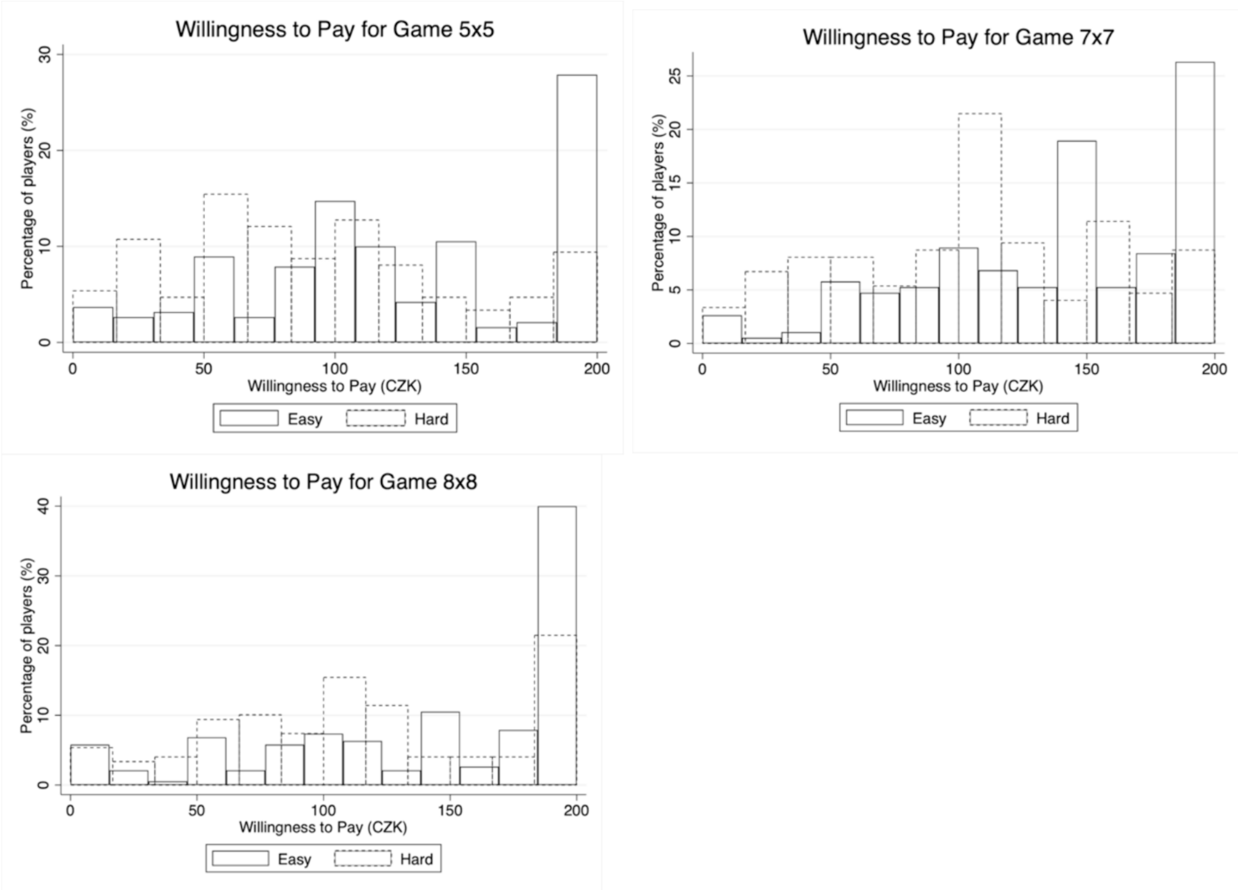

Fig. 2. Distribution of bids for games in the third round.

easy conditions. Contrary to H2c, bids by the followers did not significantly differ from leaders' bids; Wilcoxon rank-sum test for difference of the bids of leaders and followers for the 5x5 matrix: $z = -0.115$, $p = 0.987$, for the bids for the 7x7 matrix: $z = 0.796$, $p = 0.426$, and for the bids for the 8x8 game: $z = 1.010$, $p = 0.312$.

Table 6 presents how the bids are influenced by the selection of the game condition and the score from the second round. In the third round bids on games, the estimations from Models 9–14 indicated that previously experienced game condition significantly influenced the bids for all game types. For the 5x5 game (Model 9), players from the hard condition adjusted their bids downward by 26.58 units relative to those in the easy condition ($b = -26.58$, $SE=7.99$, $p < 0.001$). The overall average willingness to pay, regardless of condition, was considerably high, with a constant value of 118.49 ($SE=6.53$, $p < 0.001$). Turning to the 7x7 game (Model 11), players from the hard condition reduced their bids by 30.92 units in comparison to the easy condition participants ($b = -30.92$,

Table 6
Bids on Games – Average Group Willingness to Pay.

	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14
	Game 5x5	Game 5x5	Game 7x7	Game 7x7	Game 8x8	Game 8x8
	ME	ME	ME	ME	ME	ME
Hard	−26.58***		−30.92***		−27.94***	
	(7.99)		(7.15)		(7.55)	
Easy X Group Score		6.628***		7.606***		7.682***
		(1.321)		(1.135)		(1.328)
Easy X Group Score		6.687***		8.551***		10.67***
		(2.578)		(2.327)		(2.723)
Constant	118.49***	65.57***	135.43***	71.9***	142.06***	73.46***
	(6.53)	(10.64)	(5.28)	(9.567)	(4.34)	(11.2)
Observations	339	339	339	339	339	339
Number of groups	23	23	23	23	23	23

***p < 0.001, **p < 0.01, *p < 0.05.

SE=7.15, $p < 0.001$). The general willingness to pay for this game type was also marked, with a constant of 135.43 (*SE*=5.28, $p < 0.001$). Regarding the 8x8 game (Model 13), the hard-condition players decreased their bids by 27.94 units in relation to the easy-condition players (b = -27.94, *SE*=7.55, $p < 0.001$). This game had the highest overall willingness to pay, as indicated by a constant of 142.06 (*SE*=4.34, $p < 0.001$). Conclusively, prior game conditions are pivotal in determining players' bidding behavior, with challenging scenarios leading to consistently lower bids across various game dimensions. The findings strongly support Hypothesis H2b, suggesting that groups playing under the easy condition exhibit a higher willingness to pay than those in the hard condition.

6.2. Results of the games

As can be seen from Table 7, teams playing the 8x8 matrix ended up with the highest score and consequently the highest monetary reward from the third round (*M*=210.8 CZK, *SD*=75.9), followed by the teams playing the 5x5 matrix (*M*=164.1, *SD*=46.53) and teams playing the 7x7 matrix (*M*=163.6, *SD*=104.25). For the 5x5 game, the total net reward (reward from games minus the bids) is similar for teams from both the easy and hard conditions (teams from easy condition: *M*=163.13, *SD*=46.47; teams from hard condition: *M*=168, *SD*=48.6, $t(58)$ = -0.32, p = 0.374). In the 7x7 game, the total net reward was higher for the hard condition (*M*=184.92, *SD*=112.14) compared to the easy condition (*M*=139.86, *SD*=90.43), $t(72)$ = -1.89, p = 0.032. And the net reward is higher for the easy condition in the 8x8 game (*M*=231.5, *SD*=54.76) compared to the hard condition (*M*=188.3, *SD*=89.2), $t(67)$ = 2.5, p = 0.009. These results don't suggest that the easy-condition players would be significantly overconfident to the point that it would decrease their profits.

6.3. Final evaluations

After the third round, followers were asked to evaluate leaders' abilities over the entire experiment, and leaders were asked to self-evaluate their abilities over every experiment; these results are summarized in Table 6. Evaluation of leaders by followers is highest for the teams playing the 5x5 game, followed by the teams playing the 8x8 game and the teams playing the 7x7 game. Leader's self-evaluations are highest for the 8x8 game, followed by the leaders playing the 7x7 and 5x5 games. When comparing the evaluation of leaders by followers in the easy and hard conditions in the first and second rounds, the evaluation is 8% lower in the teams that played the hard condition (M=5.26, SD=1.8) compared to the easy condition (*M*=5.77, *SD*=1.95), $t(224) = 1.99$, $p = 0.024$, which is a decrease of effect of misattribution on evaluation compared to the second round. The difference between the leaders' self-evaluation

Table 7
Evaluation of Third Round.

	Overall				Easy condition in first two rounds				Hard condition in first two rounds			
Variable	N	Mean	Median	St. Dev.	N	Mean	Median	St. Dev.	N	Mean	Median	St. Dev.
Third Round – Overall (203 players played, 60% of all participants)												
Game Score R3	203	6	4	3.27	119	6.5	7	3.33	84	5.28	4	3.1
Team Reward R3	203	179.8	190	83.64	119	177	194	74.16	84	183.82	177	95.82
Satisfaction of results R3	203	5.92	6	2.6	119	6.03	7	2.78	84	5.75	6	2.32
Player self-evaluation R3	203	6.03	7	2.8	119	6.22	7	2.87	84	5.75	6	2.7
Team evaluation R3	203	6.53	7	2.35	119	6.82	7	2.3	84	6.13	6	2.38
Leader evaluation R3	135	6.06	6	1.8	79	6.28	6	1.77	56	5.75	6	1.8
Leader self-evaluation R3	68	6.02	6	1.96	40	6.33	6	1.59	28	5.57	5	2.35
Third round, Game 5x5 (n = 60, 30% of round 3 participants, 17.7% of total participants)												
Game Score R3 5x5	60	9.95	10	0.29	48	9.94	10	0.32	12	10	10	0
Team Reward R3 5x5	60	164.1	172.5	46.53	48	163.1	172.5	46.5	12	168	181.5	48.6
Satisfaction of results R3 5x5	60	8.15	9	1.8	48	8.15	9	1.9	12	8.17	9	1.4
Player self-evaluation R3 5x5	60	8.52	9	1.52	48	8.44	9	1.67	12	8.83	9	0.58
Team evaluation R3 5x5	60	8.57	9	1.25	48	8.54	9	1.34	12	8.67	9	0.89
Leader evaluation R3 5x5	40	6.5	7	1.71	32	6.47	7	1.78	8	6.625	7	1.41
Leader self-evaluation R3 5x5	20	6.5	6.5	1.76	16	6.75	7	1.84	4	5.5	5	1
Third round, Game 7x7 (n = 74, 36.5% of round 3 participants, 21.8% of total participants)												
Game Score R3 7x7	74	4.81	4	2.87	35	4.4	3	2.75	39	5.18	5	2.95
Team Reward R3 7x7	74	163.6	128	104.25	35	139.9	117	90.4	39	184.9	145	112.1
Satisfaction of results R3 7x7	74	5.16	5	2.33	35	4.63	4	2.39	39	5.64	6	2.2
Player self-evaluation R3 7x7	74	5.12	5	2.72	35	4.77	5	2.78	39	5.44	5	2.67
Team evaluation R3 7x7	74	5.89	6	2.26	35	5.74	6	2.24	39	6.03	6	2.3
Leader evaluation R3 7x7	49	5.98	6	1.64	23	6.13	6	1.58	26	5.85	6	1.71
Leader self-evaluation R3 7x7	25	5.52	6	2.14	12	5.58	5.5	1.24	13	5.47	6	2.78
Third round, Game 8x8 (n = 69, 34% of round 3 participants, 20 % of total participants)												
Game Score R3 8x8	69	3.83	4	1.68	36	3.94	4	1.62	33	3.7	4	1.76
Team Reward R3 8x8	69	210.8	195	75.9	36	231.5	222.5	54.76	33	188.3	190	89.2
Satisfaction of results R3 8x8	69	4.8	4	2.25	36	4.56	4	2.3	33	5.06	5	2.2
Player self-evaluation R3 8x8	69	4.83	5	2.4	36	4.7	4	2.93	33	5	6	2.44
Team evaluation R3 8x8	69	5.45	6	2.08	36	5.56	6	1.92	33	5.33	6	2.26
Leader evaluation R3 8x8	46	5.76	6	2	24	6.17	6	1.95	22	5.32	6	2
Leader self-evaluation R3 8x8	23	6.13	6	1.87	12	6.5	6.5	1.38	11	5.73	5	2.28

for the easy (M=5.77, SD=1.95) and hard conditions in the first two rounds (M=5.2, SD=2.1) is again insignificant, as after the second round, $t(111) = 1.47$, $p = 0.073$. However, it is important to acknowledge that the teams were not assigned to play either of the games in the third round randomly, and the final evaluation of the leader and leader's self-evaluation are arguably a mix of the score in the first two rounds and the result from the third round.

6.4. Self-assessment manikins, self-efficacy, and sense of power

After the first round, we found that players who played in the easy condition were happier (M=4.98, SD=1.83) compared to those playing in the hard condition (M=4.01, SD=1.93), $t(337) = -4.77$, $p < 0.001$. Also, after the first round, players in the easy condition had a higher intensity of emotions (M=3.58, SD=2.47) than the hard condition players (M=5.027, SD=2.29), $t(337) = -5.51$, $p < 0.001$. After the first round, players in the easy condition also felt more powerful (M=5.42, SD=2.1) than the hard condition players (M=4.75, SD=1.92), $t(337) = 3.003$, $p < 0.001$. After the second round, the easy-condition players were happier (M=5.24, SD=1.86) than the hard-condition players (M=5.57, SD=1.77), $t(337) = -1.67$, $p = 0.047$, but at this time, there was no significant difference neither in the intensity of emotions, $t(337) = -0.385$, $p = 0.350$, nor in the sense of power $t(337) = -0.560$, $p = 0.562$. In terms of the Sense of power, the players felt with regard to being either the follower or the leader, there were no differences in answers, neither after the first round, $t(337) = -0.017$, $p = 0.490$, nor in the second round, $t(337) = -0.45$, $p = 0.328$. Complete results of the self-assessment manikins according to both the condition and the position in the team, including the results after the third round, are reported in Table OA1 in the online appendix.

Neither the self-efficacy nor sense of power questionnaires produced any statistically significant differences between the answers from easy and hard-condition players, with one exception:in the sense of power questionnaire, easy-condition players tended to agree more strongly (M=2.42, SD=1.58) than the hard-condition players (M=2.01, SD=1.53) with the statement "Even if I voice them, my views have little sway", $t(337) = 2.35$, $p = 0.020$. Complete results are reported in the online appendix (Tables OA2 and OA3).

6.5. Exploratory analysis of gender effect

The average personal score from the first two rounds was influenced by a gender effect. Over the first two rounds, females playing in the easy condition over the first two rounds scored on average 6.2% less (M=8.16, SD=2.21) during each round than males (M=8.70, SD=2.32), $t(188) = 1.63$, $p = 0.052$. In the hard condition, females scored on average 21% less (M=2.89, SD=1.45) each round than males (M=3.67, SD=1.71), $t(147) = 2.9$, $p = 0.002$. After the second round, females playing in the hard condition rated their leaders less harshly than males – they rated their leaders 23.3% lower (M=5.21, SD=2) than females playing in the easy condition (M=6.78, SD=2.15), $t(106) = 3.85$, $p < 0.001$. Males playing in the hard condition rated their leaders 29.5% lower (M=6.21, SD=2.2) than males in the easy condition (M=4.70, SD=2.1), $t(116) = 4.63$, $p < 0.001$. Therefore, females playing in the easy condition rated their leaders 9.5% higher (M=6.78, SD=2.15) than males (M=6.21, SD=2.15), $t(124) = -1.47$, $p = 0.070$, while females playing in the hard condition rated their leaders 18.7% higher (M=5.21, SD=2) than males (M=4.38, SD=5.2), $t(98) = -2.0$, $p = 0.024$. On the other hand, female leaders did not evaluate themselves significantly differently between the easy (M=5.85, SD=2.13) and hard conditions (M=5.26, SD=2), $t(43) = 0.93$, $p = 0.180$, and neither did male leaders: easy (M=5.55, SD=2.14); hard (M=5.4, SD=2.09), $t(66) = 0.3$, $p = 0.380$. Female leaders playing in the hard condition were 23.4% less satisfied with the work of other team members (M=6.10, SD=7.96) compared to the females playing in the easy condition (M=7.96, SD=1.23), $t(43) = 3.58$, $p < 0.001$. Male leaders playing in the hard condition rated the work of other team members 37% lower (M=4.63, SD=2.67) than male leaders playing in the easy condition (M=7.34, SD=2.47), $t(66) = 4.32$, $p < 0.001$. Overall, female leaders playing in the hard condition were 32% more satisfied with the work of other team members (M=6.12, SD=2.18) compared to the male leaders (M=4.63, SD=2.68), $t(47) = -2.01$, $p = 0.025$. For the easy condition, there was no significant difference in the leaders' satisfaction with the work of other team members between female leaders (M=7.96, SD=2.47) and male leaders (M=7.34, SD=2.47), $t(62) = -1.7$, $p = 0.120$.

6.6. Robustness analysis – Income effect

In our analysis, we acknowledged that earnings varied significantly between the conditions in the first two rounds. Consequently, we conducted a supplementary robustness analysis to address the potential confounding influence of earnings in tandem with task difficulty. To ensure that our results are not driven exclusively by differences in earnings, we introduced earnings as a control variable in our regression models. Specifically, we adjusted our models to incorporate the sum of earnings from the first two rounds (EarningsR1R2) as a covariate. This adjustment aims to distinguish the unique impact of task difficulty from potential earnings effects.

Results of this robustness test for the evaluation of the leader and self-valuation after the second round are detailed in Table OA4 in the online appendix. As evident from the table, when the models using the game condition as a treatment effect (models A1, A2) incorporate the earnings variable, the coefficient is positive and statistically significant ($p < 0.001$). This suggests an income effect, where higher earnings have a positive, albeit modest, influence on both the evaluation of the leader by followers and the leader's self-evaluation. Notably, the treatment condition's estimated coefficients in Table 4 (Models 1, 5) display only slight reductions in magnitude, and their original statistical significance remains largely intact, even after accounting for earnings. Thus, we believe our conclusions regarding the treatment effect of the game condition on leader evaluations remain robust against the income effect.

Similarly, the income effect robustness results for the bids on games in the third round are outlined in Table OA5 in the online appendix. Here, introducing earnings from the first two rounds alongside the treatment condition yields a statistically significant coefficient ($p < 0.001$). However, the treatment condition's effect remains statistically significant across all three games (Game 5x5, 7x7, 8x8 as represented by models A7, A8, and A9). The magnitude and direction are consistent with our expectations, suggesting that teams originally assigned to the hard condition in the first two rounds bid, on average, less for games in the third round compared to teams from the easy condition. In summary, while earnings do play a role, our primary conclusions concerning the treatment effect of the game condition hold strong and are further solidified when controlling for earnings.

7. Discussion

This study showed that random situational factors (easy or hard conditions of a game) impacted the evaluation of leaders' and followers' competencies and identified the consequences of this incorrect attribution of merit. Although the situational factor and randomly worse outcomes did not affect the leaders' self-evaluation (they assessed themselves equally well across conditions), team members were rated worse by leaders in difficult conditions. Also, team members in the hard condition rated their contribution to team performance worse than the members playing the easy one. We have named this dynamic – people are more generous in praising the leader's success than with their own performance, even if the success is completely random – the leadership fallacy.

The replicated study, Weber et al. (2001), excluded leaders' self-ratings in their analysis due to the low number of leaders; however, they suggested the mechanism of leaders' attribution is the same as that of followers. Our results contradict this conclusion; leaders are not influenced in their self-assessment by the conditions of a task or their team's actual performance. Our result supports the self-serving tendency of leaders (Meindl & Ehrlich, 1987). Indeed, the asymmetric attributions of outcomes based on a person's position in the team can result in systematic mistakes in assessment (Houdek et al., 2024): promotions not reflecting actual performance or incorrect apportioning of blame for bad results to perverse talent management (Liu & de Rond, 2016).

If poor performance is attributed to team members, good performance to the leaders and the self-evaluation of the leaders is good in all conditions, the dynamics may lead to rewarding executives based on luck (Houdek, 2017). This illusion of success may harm organizations in the form of a lack of understanding about the real competencies of its members. Mistakes that are not rectified can accumulate and lead to companies' failures (Chen et al., 2015). The findings from Cavalcantii et al. (2023) further complicate this issue, showing that people may follow leaders' directions even when those leaders lack relevant information or talent, highlighting the potential for leadership influence based on irrelevant attributes. Explaining the asymmetry in attribution based on team position has some important implications for the allocation of responsibility and adequate remuneration of executives. If companies consistently attribute their successes or losses to the executives' skills without considering external market conditions, they may also fail to recognize potential weaknesses or areas for improvement. They can reward faulty learning of their leaders (by selection neglect, they will imitate random survivors; Barron et al., 2019; Jehiel, 2018), excessive optimism in one's abilities, and lack of benchmarking, i.e., underestimating competitors. Reducing the attribution biases may be appropriate to corporate governance (Bertrand & Mullainathan, 2001).

Evaluation of leaders based on conditions is in congruence with the romanticizing tendencies of followers (Bligh et al., 2011). The more successful a team was, the better its leader was rated, although the leader had no influence on the team's performance. Leaders playing in easy condition were rated far better by the team members, which is a classic demonstration of attribution error in group assessment (Hewstone, 1990). The effect remains even after the second and third rounds of the experiment, so the results are not just the effect of an atypical, one-round situation but point out the prevailing tendency in assessment. An alternative explanation may be that the leader had the opportunity to act (give advice) while other team members did not, which could induce an (incorrect) impression that the behavior of a leader is the only possible cause for the outcome (Kahneman & Miller, 1986).

In contrast to the study by Weber et al. (2001), in our study, participants could improve their performance over the course of the experiment by learning. In the second round of the experiment, the performance of the teams improved in both easy and hard conditions. Although the teams playing in hard condition improved more than the teams playing in easy condition, their leaders were still rated worse than the leaders playing in an easy condition. Thus, it seems that identifying improvement in difficult conditions is more demanding. As mentioned above, team members playing in the hard condition rated their contribution to team performance worse than the members from the easy one. Even if workers do the maximum in a given condition, they evaluate their own abilities based on (unattainable) results. In the words of the discounting principle theory (Rosenfield & Stephan, 1977), people don't employ sufficient mental schema for causes of an outcome and overlook important features of a situation – the evaluation should correspond to an improvement in performance, not to an initial (random) disadvantage that the leader had no control over. A poor reflection of one's possible competence in a situation can be an important predictor of burnout (Bianchi & Schonfeld, 2016).

Our results are also consistent with the "beginner bubble" hypothesis of overconfidence occurring within the experience of just a few trials (Sanchez & Dunning, 2018). Participants gained experience and feedback on the tasks in the first two rounds. Players (leaders as well as followers) playing in easy condition were conditioned by this experience, which resulted in a boost of confidence. Each success in the task served as evidence of the team's ability for leaders who decided to continue in the third round. People tend to under-sample failures and focus on survivors (Denrell, 2003). Because success was easily recalled from memory in the easy task group, the choice to continue in the game was substantiated by the "lessons learned." Given the structure of our experiment, engaging in the next round was beneficial for the team. Different task conditions could lead to different outcomes. The underestimation of randomness can have fatal consequences for organizational learning and project and strategic management (Liu & de Rond, 2016). A possible alternative explanation for the overconfidence of easy-task teams may be that they had a worse idea of the difficulty of the most difficult game (than hard-task teams), which in turn led to a greater willingness to pay for the most difficult game.

Manipulations by the task difficulty and assignment of leaders did not affect the constructs of self-efficacy and sense of power. Thus, these observed constructs did not moderate the observed behavior as we predicted. Success in the easy task did not evoke differences in the sense of power, probably because of task simplicity and the laboratory environment. Although individuals assigned to the leader's position had control over their decision (Anderson et al., 2012) to continue in the game, this difference in team roles was not perceived so markedly as to induce changes in perceived psychological states even though the experimental manipulation influenced observed behavior (Galinsky et al., 2003).

In the exploratory analysis, we examined some differences in assessment based on gender. We found that women were more forgiving in assessing leaders for both hard and easy conditions than men. This difference was even more pronounced in the difficult condition. One of the possible explanations is the women's lower ability to perform the task, which could lead them to blame themselves for failure (Andrews, 1987), taking the blame off the leader. An alternative explanation could be women's tendency to be more lenient in assessment in general (Shore et al., 1997). However, female leaders, as well as male leaders, did not adjust their self-assessment across conditions. Thus, women were subject to the effect of misattribution in self-evaluation in the same way as men.

More work must be done to understand whether examined biases exist in other contexts. For example, it is necessary to determine how dispositional factors affect leader evaluation and team self-evaluation in a real environment or how personality characteristics can mediate the evaluation and consequent behavior. Future research could also investigate why a significant improvement in a difficult environment does not lead to a better evaluation of the leader. Research on how a team's successes or failures shape beliefs (and how they are cognitively processed) that influence cooperation between leaders and followers would also be a fruitful avenue.

Funding and acknowledgments

This project was supported by the Faculty of Business Administration, Prague University of Economics and Business (an internal grant IG310029); P. H. acknowledges funding from the Czech Science Foundation Project No. 22-29520S and UNCE 24/SSH/018. We would like to thank Marek Hudík, Lucie Vrbová, Marek Vranka, Jakub Procházka, and Štěpán Bahník for their inspiring comments, as well as to the attendees of the EURAM 2022 Doctoral Colloquium in Dublin. Our appreciation also goes to our research assistants, Marek Pour and Marketa Sýkorová, for data collection and to Jan Vávra for programming the experiment. Finally, we also thank the reviewers and the editor, Carlos Alós-Ferrer, whose helpful comments and ideas greatly improved the quality of our manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have share the link to our data in the Data Availability Statement.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.joep.2024.102753.

References

Ahle, T. D. (2012/2020). *Thomasahle/numberlink* [Go]. https://github.com/thomasahle/numberlink (Original work published 2012).

Anderson, C., John, O. P., & Keltner, D. (2012). The Personal Sense of Power: The Personal Sense of Power. *Journal of Personality, 80*(2), 313–344. https://doi.org/10.1111/j.1467-6494.2011.00734.x

Andrews, P. H. (1987). Gender Differences in Persuasive Communication and Attribution of Success and Failure. *Human Communication Research, 13*(3), 372–385. https://doi.org/10.1111/j.1468-2958.1987.tb00110.x

Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly, 20*(5), 764–784. https://doi.org/10.1016/j.leaqua.2009.06.006

Barber, B. M., & Odean, T. (2001). Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics*, *116*(1), 261–292. doi: 10.1162/003355301556400.

Barron, K., Huck, S., & Jehiel, P. (2019). *Everyday econometricians: Selection neglect and overoptimism when learning from others* (No. SP II 2019-301). WZB Discussion Paper.

Bertrand, M., & Mullainathan, S. (2001). Are CEOs Rewarded for Luck? The Ones Without Principals Are. *The Quarterly Journal of Economics, 116*(3), 901–932. https://doi.org/10.1162/00335530152466269

Bianchi, R., & Schonfeld, I. S. (2016). Burnout is associated with a depressive cognitive style. *Personality and Individual Differences, 100*, 1–5. https://doi.org/10.1016/j.paid.2016.01.008

Bligh, M. C., Kohles, J. C., & Pillai, R. (2011). Romancing leadership: Past, present, and future. *The Leadership Quarterly, 22*(6), 1058–1077. https://doi.org/10.1016/j.leaqua.2011.09.003

Braun, P. A., & Yaniv, I. (1992). A case study of expert judgment: Economists' probabilities versus base-rate model forecasts. *Journal of Behavioral Decision Making, 5*(3), 217–231. https://doi.org/10.1002/bdm.3960050306

Camerer, C., & Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review, 89*(1), 306–318. https://doi.org/10.1257/aer.89.1.306

Cavalcanti, C., Grossman, P. J., & Khalil, E. L. (2023). Leadership heuristic. *Journal of Economic Psychology, 98*, Article 102661. https://doi.org/10.1016/j.joep.2023.102661

Chen, G., Crossland, C., & Luo, S. (2015). Making the same mistake all over again: CEO overconfidence and corporate resistance to corrective feedback. *Strategic Management Journal, 36*(10), 1513–1535. https://doi.org/10.1002/smj.2291

Cheung, F., & Lucas, R. E. (2014). Assessing the validity of single-item life satisfaction measures: Results from three large samples. *Quality of Life Research, 23*(10), 2809–2818. https://doi.org/10.1007/s11136-014-0726-4

Coutts, A., Gerhards, L., & Murad, Z. (2020). *Who to blame? Self-serving attribution bias with multi-dimensional uncertainty*. Working Paper.

Denrell, J. (2003). Vicarious Learning, Undersampling of Failure, and the Myths of Management. *Organization Science, 14*(3), 227–243. https://doi.org/10.1287/orsc.14.2.227.15164

Dillon, R. L., & Tinsley, C. H. (2008). How Near-Misses Influence Decision Making Under Risk: A Missed Opportunity for Learning. *Management Science, 54*(8), 1425–1440. https://doi.org/10.1287/mnsc.1080.0869

Dobbins, G. H., & Russell, J. M. (1986). Self-Serving Biases in Leadership: A Laboratory Experiment. *Journal of Management, 12*(4), 475–483. https://doi.org/10.1177/014920638601200403

Elsesser, K. M. (2016). Gender Bias Against Female Leaders: A Review. In M. L. Connerley, & J. Wu (Eds.), *Handbook on Well-Being of Working Women* (pp. 161–173). Netherlands: Springer. https://doi.org/10.1007/978-94-017-9897-6_10.

Enke, B. (2020). What you see is all there is. *The Quarterly Journal of Economics, 135*(3), 1363–1398. https://doi.org/10.1093/qje/qjaa012

Ensari, N., & Murphy, S. E. (2003). Cross-cultural variations in leadership perceptions and attribution of charisma to the leader. *Organizational Behavior and Human Decision Processes, 92*(1), 52–66. https://doi.org/10.1016/S0749-5978(03)00066-9

Erkal, N., Gangadharan, L., & Koh, B. H. (2022). By chance or by choice? Biased attribution of others' outcomes when social preferences matter. *Experimental Economics, 25*(2), 413–443. https://doi.org/10.2139/ssrn.3251105

Fabricius, G., & Büttgen, M. (2015). Project managers' overconfidence: How is risk reflected in anticipated project success? *Business Research, 8*(2), 239–263. https://doi.org/10.1007/s40685-015-0022-3

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From Power to Action. *Journal of Personality and Social Psychology, 85*(3), 453–466. https://doi.org/10.1037/0022-3514.85.3.453

Garicano, L., & Rayo, L. (2016). Why Organizations Fail: Models and Cases. *Journal of Economic Literature, 54*(1), 137–192. https://doi.org/10.1257/jel.54.1.137

Gentry, W. A., & Sparks, T. E. (2012). A convergence/divergence perspective of leadership competencies managers believe are most important for success in organizations: A cross-cultural multilevel analysis of 40 countries. *Journal of Business and Psychology, 27*(1), 15–30. https://doi.org/10.1007/s10869-011-9212-y

Grinblatt, M., & Keloharju, M. (2009). Sensation Seeking, Overconfidence, and Trading Activity. *The Journal of Finance, 64*(2), 549–578. https://doi.org/10.1111/j.1540-6261.2009.01443.x

Hambrick, D. C. (2007). Upper Echelons Theory: An Update. *Academy of Management Review, 32*(2), 334–343. https://doi.org/10.5465/amr.2007.24345254

Harrison, J. R., & March, J. G. (1984). Decision Making and Postdecision Surprises. *Administrative Science Quarterly*, 29(1), 26–42. doi: 10.2307/2393078.

Heidhues, P., Kőszegi, B., & Strack, P. (2018). Unrealistic expectations and misguided learning. *Econometrica, 86*(4), 1159–1214. https://doi.org/10.3982/ECTA14084

Hestermann, N., & Le Yaouanq, Y. (2021). Experimentation with self-serving attribution biases. *American Economic Journal: Microeconomics, 13*(3), 198–237. https://doi.org/10.1257/mic.20180326

Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology, 20*(4), 311–335. https://doi.org/10.1002/ejsp.242020040

Houdek, P. (2017). Rewards for falling off a horse. *Organizational Dynamics, 46*(3), 189–194. https://doi.org/10.1016/j.orgdyn.2017.05.004

Houdek, P., Bahník, Š., Vrbová, L., & Hájek, J. (2024). Perspective-Taking Does not Reduce Victim Blaming in Work-related Situations. *Deviant Behavior*. https://doi.org/10.1080/01639625.2024.2354349

James, D. (2007). Stability of risk preference parameter estimates within the Becker-DeGroot-Marschak procedure. *Experimental Economics, 10*(2), 123–141. https://doi.org/10.1007/s10683-006-9136-y

Jehiel, P. (2018). Investment strategy and selection bias: An equilibrium perspective on overoptimism. *American Economic Review, 108*(6), 1582–1597. https://doi.org/10.1257/aer.20161696

Jones, E. E., & Davis, K. E. (1965). From Acts To Dispositions The Attribution Process In Person Perception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2, pp. 219–266). Academic Press. https://doi.org/10.1016/S0065-2601(08)60107-0.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological review, 93*(2), 136. https://doi.org/10.1037/0033-295X.93.2.136

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. A. (2006). Would You Be Happier If You Were Richer? *A Focusing Illusion. Science, 312*(5782), 1908–1910. https://doi.org/10.1126/science.1129688

Kassas, B., Palma, M. A., & Porter, M. (2022). Happy to take some risk: Estimating the effect of induced emotions on risk preferences. *Journal of Economic Psychology, 91*, Article 102527. https://doi.org/10.1016/j.joep.2022.102527

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265. https://doi.org/10.1037/0033-295X.110.2.265

Khalil, E. L. (2022). The information inelasticity of habits: Kahneman's bounded rationality or Simon's procedural rationality? *Synthese, 200*(4), 343. https://doi.org/10.1007/s11229-022-03815-7

Kirkpatrick, S. A., & Locke, E. A. (1996). Direct and indirect effects of three core charismatic leadership components on performance and attitudes. *Journal of Applied Psychology, 81*(1), 36. https://doi.org/10.1037/0021-9010.81.1.36

Köhler, T., & Cortina, J. M. (2019). Play It Again, Sam! An Analysis of Constructive Replication in the Organizational Sciences. *Journal of Management, 0149206319843985*. https://doi.org/10.1177/0149206319843985

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*(2), 311–328. https://doi.org/10.1037/0022-3514.32.2.311

Lassiter, G. D., Geers, A. L., Munhall, P. J., Ploutz-Snyder, R. J., & Breitenbecher, D. L. (2002). Illusory causation: Why it occurs. *Psychological Science, 13*(4), 299–305. https://doi.org/10.1111/j.0956-7976.2002.x

Libby, R., & Rennekamp, K. (2012). Self-Serving Attribution Bias, Overconfidence, and the Issuance of Management Forecasts: Overconfidence and forecast issuance. *Journal of Accounting Research, 50*(1), 197–231. https://doi.org/10.1111/j.1475-679X.2011.00430.x

Liu, C., & de Rond, M. (2016). Good Night, and Good Luck: Perspectives on Luck in Management Scholarship. *Academy of Management Annals, 10*(1), 409–451. https://doi.org/10.5465/19416520.2016.1120971

Loerakker, B., & van Winden, F. (2017). Emotional Leadership in an Intergroup Conflict Game Experiment. *Journal of Economic Psychology, 63*, 143–167. https://doi.org/10.1016/j.joep.2017.03.009

Lopez, E. S., & Ensari, N. (2014). The Effects of Leadership Style, Organizational Outcome, and Gender on Attributional Bias Toward Leaders. *Journal of Leadership Studies, 8*(2), 19–37. https://doi.org/10.1002/jls.21326

Lovell, C. D., & Kosten, L. A. (2000). Skills, Knowledge, and Personal Traits Necessary for Success as a Student Affairs Administrator: A Meta-Analysis of Thirty Years of Research. *NASPA Journal, 37*(4), 535–572. https://doi.org/10.2202/1949-6605.1118

Lunenburg, F. C. (2012). Power and leadership: An influence process. *International Journal of Management, Business, and Administration, 15*(1), 1–9.

Meindl, J. R., & Ehrlich, S. B. (1987). The Romance of Leadership and The Evaluation of Organizational Performance. *Academy of Management Journal, 30*(1), 91–109. https://doi.org/10.5465/255897

Meissner, P., & Wulf, T. (2017). The effect of cognitive diversity on the illusion of control bias in strategic decisions: An experimental investigation. *European Management Journal, 35*(4), 430–439. https://doi.org/10.1016/j.emj.2016.12.004

Milgrom, P., & Roberts, J. (1988). An economic approach to influence activities in organizations. *American Journal of Sociology, 94*, S154–S179. https://doi.org/10.1086/228945

Nasir Zadeh, F., Zangi Abadi, M., & Molla Imeny, V. (2020). The application of the correspondent inference theory (Attribution theory) in the behavior of investors with a quasi-empirical approach. *Journal of Accounting, Accountability and Society Interests, 10*(3), 117–134. https://doi.org/10.22051/ijar.2020.26858.1521

Pluchino, A., Biondo, A. E., & Rapisarda, A. (2018). Talent versus luck: The role of randomness in success and failure. *Advances in Complex Systems, 21*(03n04), 1850014. https://doi.org/10.1142/S0219525918500145

Porter, L. W., & McLaughlin, G. B. (2006). Leadership and the organizational context: Like the weather? *The Leadership Quarterly, 17*(6), 559–576. https://doi.org/10.1016/j.leaqua.2006.10.002

Quan, P., Feng, C., & Yang, Q. (2011). Lacking control: Situation valence, effects of perceived control, and illusory pattern perception. *Social Behavior and Personality: An International Journal, 39*(10), 1413–1416. https://doi.org/10.2224/sbp.2011.39.10.1413

Ridgell, S. D., & Lounsbury, J. W. (2004). Predicting academic success: General intelligence,“ Big Five” personality traits, and work drive. *College Student Journal, 38*(4), 607–619. https://link.gale.com/apps/doc/A126386901/AONE?u=anon~9bd565e0&sid=googleScholar&xid=fe38c260.

Ronay, R., Oostrom, J. K., Lehmann-Willenbrock, N., Mayoral, S., & Rusch, H. (2019). Playing the trump card: Why we select overconfident leaders and why it matters. *The Leadership Quarterly, 30*(6), Article 101316. https://doi.org/10.1016/j.leaqua.2019.101316

Ronis, D. L., & Yates, J. F. (1987). Components of probability judgment accuracy: Individual consistency and effects of subject matter and assessment method. *Organizational Behavior and Human Decision Processes, 40*(2), 193–218. https://doi.org/10.1016/0749-5978(87)90012-4

Rosenfield, D., & Stephan, W. G. (1977). When discounting fails: An unexpected finding. *Memory & Cognition, 5*(1), 97–102. https://doi.org/10.3758/BF03209199

Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of personality and social psychology, 35*(7), 485. https://doi.org/10.1037/0022-3514.35.7.485

Salancik, G. R., & Meindl, J. R. (1984). Corporate Attributions as Strategic Illusions of Management Control. *Administrative Science Quarterly*, *29*(2), 238–254. doi: 10.2307/2393176.

Sanchez, C., & Dunning, D. (2018). Overconfidence among beginners: Is a little learning a dangerous thing? *Journal of Personality and Social Psychology, 114*(1), 10–28. https://doi.org/10.1037/pspa0000102

Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, W. K. (2004). Overconfidence and the big five. *Journal of Research in Personality, 38*(5), 473–480. https://doi.org/10.1016/j.jrp.2003.09.010

Schilke, O., Levine, S. S., Kacperczyk, O., & Zucker, L. G. (2019). Call for Papers-Special Issue on Experiments in Organizational Theory. *Organization Science*, *30*(1), 232–234. doi: 10.1287/orsc.2018.1257.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. *Measures in Health Psychology: A User's Portfolio*. *Causal and Control Beliefs*, (1), 35–37. https://doi.org/10.4236/jss.2014.211021

Shore, T., Tashchian, A., & Adams, J. (1997). The Role of Gender in a Developmental Assessment Center. *Journal of Social Behavior & Personality, 12*(5), 191–203. https://doi.org/10.1037/0021-9010.91.3.555

Simon, J. G., & Feather, N. T. (1973). Causal attributions for success and failure at university examinations. *Journal of Educational Psychology, 64*(1), 46–56. https://doi.org/10.1037/h0034083

Standing, C., Guilfoyle, A., Lin, C., & Love, P. E. D. (2006). The attribution of success and failure in IT projects. *Industrial Management & Data Systems, 106*(8), 1148–1165. https://doi.org/10.1108/02635570610710809

Swift, S. A., Moore, D. A., Sharek, Z. S., & Gino, F. (2013). Inflated Applicants: Attribution Errors in Performance Evaluation by Professionals. *PLOS ONE, 8*(7), e69258.

Vosgerau, J. (2010). How prevalent is wishful thinking? Misattribution of arousal causes optimism and pessimism in subjective probabilities. *Journal of Experimental Psychology: General, 139*(1), 32–48. https://doi.org/10.1037/a0018144

Wageman, R., & Richard Hackman, J. (2007). Asking the Right Questions About Leadership: Discussion and Conclusions. *American Psychologist, 62*. https://doi.org/10.1037/0003-066X.62.1.43

Weber, R., Camerer, C., Rottenstreich, Y., & Knez, M. (2001). The Illusion of Leadership: Misattribution of Cause in Coordination Games. *Organization Science*, *12*(5), 582–598. doi: 10.1287/orsc.12.5.582.10090.

Whitson, J. A., & Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science, 322*(5898), 115–117. https://doi.org/10.1126/science.1159845

Williams, D. (2021). Socially adaptive belief. *Mind & Language, 36*(3), 333–354. https://doi.org/10.1111/mila.12294

Yazici, H. J. (2011). Significance of organizational culture in perceived project and business performance. *Engineering Management Journal, 23*(2), 20–29. https://doi.org/10.1080/10429247.2011.11431892