

QUITTING WHEN THE GOING GETS TOUGH: A DOWNSIDE OF HIGH PERFORMANCE EXPECTATIONS

HENGCHEN DAI

University of California, Los Angeles

BERKELEY J. DIETVORST

University of Chicago

BRADFORD TUCKFIELD

American Express GBT

KATHERINE L. MILKMAN

MAURICE E. SCHWEITZER

University of Pennsylvania

High performance expectations often improve performance. When individuals with high external performance expectations encounter early setbacks, however, they face impression management concerns and the prospect of embarrassment. As a result, when the going gets tough, individuals facing high external expectations may be less likely to persist than people facing low external expectations. In a field study of 328,515 men's professional tennis matches (Study 1), we employ a regression discontinuity design to demonstrate that, after losing the first set of a match, players who are expected to win (favorites) are significantly more likely to quit than players who are expected to lose (underdogs). We replicate this pattern of results in a laboratory experiment (Study 2) and provide evidence for our proposed mechanism: compared to individuals facing low external expectations, those facing high expectations are more easily embarrassed by poor performance and consequently less persistent following early setbacks.

Executives, managers, and employees constantly face decisions about whether to persist in their current endeavor or redirect their efforts. For example, entrepreneurs regularly make decisions about whether to persist with or abandon their ventures (Gimeno, Folta, Cooper, & Woo, 1997), project managers decide to persevere with or terminate their projects (Green, Welsh, & Dehler, 2003), and employees make frequent

decisions regarding whether to continue with a current task or redirect their efforts (Conlon, 1980). The decision to persist or quit can have profound implications. Persisting in a failing venture can be ruinous, but quitting prematurely may mean forgoing opportunities to reap substantial rewards (e.g., Brooks & Schweitzer, 2011; March, 1991; Weber & Camerer, 1998; Zhang, Allon, & Van Mieghem, 2017). In this work, we investigate the decision to persist and focus on the consequences of an important antecedent: external performance expectations.

We pursue many of our most consequential endeavors in the face of external performance expectations (e.g., from our supervisors, coworkers, clients, family, friends, and the media). Existing research reveals that, compared to people who face low external performance expectations, those who face high expectations expend greater effort, exhibit greater persistence, and achieve better outcomes (e.g., Eden, 1990; McNatt, 2000). In this work, we demonstrate that the relationship between performance expectations and persistence is more complicated than prior research

We thank associate editor Gerben van der Vegt and three anonymous reviewers for their extremely thorough and constructive guidance throughout the review process. We thank Markus Baer, Hillary Anger Elfenbein, Adam Grant, Samir Nurmohamed, Bradley Staats, and Dennis Zhang, as well as members of the Operations, Information, and Decisions Department at Wharton for their insightful comments on earlier drafts of this paper. We thank research assistants at the Wharton Behavioral Lab and Olin Business School for their help collecting data.

This research received financial support from the Wharton Behavioral Lab, Olin Business School, the Wharton Sports Business Initiative, and the Zicklin Center for Business Ethics Research.

suggests. We extend existing theory by proposing and empirically demonstrating that high external expectations can *harm* persistence under some circumstances when people have impression management motives (i.e., when they aim to maintain a positive public image; Leary & Kowalski, 1990). Specifically, we show that, after exhibiting poor initial performance on a task, individuals who face high external expectations feel more embarrassed about violating public expectations and in turn are less persistent than individuals who face low external expectations.

EXPECTATIONS, PERFORMANCE, AND PERSISTENCE

In general, performance is jointly influenced by a host of individual characteristics (e.g., motivation and ability) and situational factors (e.g., task difficulty, luck, collaborators' or competitors' motivation and ability). Even strong and highly motivated performers will occasionally perform poorly. We define "poor initial performance" as performance in an early stage of an endeavor that creates a substantial obstacle for success. After experiencing poor performance in the initial stage of a task or pursuit, individuals can either persist or shift their efforts elsewhere. We investigate how external performance expectations influence this decision.

External performance expectations (expectations held by observers about a target individual) may result from prior experiences (Mishina, Dykes, Block, & Pollock, 2010), rankings (Luca & Smith, 2013; Pope, 2009), reputation (Jensen, Kim, & Kim, 2012; Petkova, Wadhwa, Yao, & Jain, 2014), or stereotypes about a group (Steele & Aronson, 1995). Research on the "Pygmalion effect" and the "Golem effect" (e.g., Babad, Inbar, & Rosenthal, 1982; Davidson & Eden, 2000; Eden, 1990; McNatt, 2000; Rosenthal & Jacobson, 1968) has found that individuals who face higher performance expectations exhibit greater effort, persistence, and performance than others. This is in part because high external expectations lead observers (e.g., managers, mentors, peers, or the media) to interact with performers differently, causing performers to experience high self-expectations (i.e., improved beliefs about their ability to perform effectively on a given task), which in turn motivates performers to exert greater effort and ultimately improve their performance (Eden, 1990).

Prior work suggests that high performance expectations should increase persistence and effort even following poor initial performance. High self-expectations enhance people's forecasts of their ability to redress an initial failure (Eden, 1990),

which are positively associated with their willingness to sustain or even heighten effort following setbacks (Bandura, 1982, 1997). People with high self-expectations are more motivated when they are slightly behind in a competition (Berger & Pope, 2011), more likely to engage in job-search activities following a job loss (Eden & Aviram, 1993), and more likely to persist even in a failing enterprise (Whyte & Saks, 2007; Whyte, Saks, & Hook, 1997). Taken together, past research suggests that, just as high external expectations elevate self-expectations (Davidson & Eden, 2000; Eden, 1990), individuals who face high external expectations should be more likely than others to sustain (or even increase) their effort following poor initial performance.

In our investigation, we challenge this assertion and demonstrate that high external performance expectations can harm persistence under adverse circumstances. We identify impression management concerns as an important and previously neglected factor that can cause people with high external performance expectations to reduce persistence in the face of adversity.

THE ROLE OF IMPRESSION MANAGEMENT CONCERNS

External Expectations and Impression Management Concerns

Individuals care deeply about maintaining a positive public image and strive to meet others' expectations to manage impressions (see Leary & Kowalski, 1990, for a review). External performance expectations (which we will refer to as "expectations" throughout the remainder of this paper for brevity) serve as reference points. Individuals expect others to evaluate their performance—and individuals evaluate their own progress—with respect to these reference points (Carver & Scheier, 1990; Leary & Kowalski, 1990). As a result, compared to individuals who face low expectations, individuals who face high expectations subject their performance outcomes to stricter standards and experience greater pressure to perform well (Gibson, Sachau, Doll, & Shumate, 2002; Lount, Pettit, & Doyle, 2017; Pettit, Sivanathan, Gladstone, & Marr, 2013). Importantly, the same objective level of performance may either conform to expectations (e.g., when an individual facing low expectations performs poorly) or violate expectations (e.g., when an individual facing high expectations performs poorly).

Poor performances can trigger impression management concerns when they occur in public (Leary

& Kowalski, 1990). We anticipate that expectations will amplify this relationship. Specifically, we predict that, compared to individuals facing low expectations, individuals facing high expectations experience greater impression management concerns after exhibiting poor performance. This prediction is consistent with the finding that, when expectations are salient and performance is publicly visible, competitors who are expected to win anticipate greater psychic costs from losing than competitors who are expected to lose (Chen, Ham, & Lim, 2011). Gibson et al. (2002) found that observers hold similar beliefs; following poor performance, observers expect individuals who face high expectations to be more concerned about impression management than individuals who face low expectations.

Performance that falls short of expectations often results in embarrassment. "Embarrassment" is a negative, self-conscious emotion triggered by undesirable events that threaten a desired public image (Miller, 1992; Modigliani, 1971; Tangney, Miller, Flicker, & Barlow, 1996). Embarrassment occurs in the company of others and is characterized by concerns about violating external standards (Miller & Leary, 1992). We propose that, when initial performance is poor, compared to individuals who face low expectations, individuals who face high expectations will be more concerned about their public image and experience greater embarrassment.

Reducing Persistence as an Impression Management Strategy

The most direct way for individuals to maintain their public image after initially falling short of expectations is to persist, improve performance, and meet expectations. However, persisting after an initial failure can be costly: individuals need to remain in an embarrassing situation and risk further, and perhaps complete, failure. Thus, rather than simply persisting, individuals often engage in impression management strategies when facing embarrassing circumstances (Leary & Kowalski, 1990).

We propose that, after an initial setback, individuals may employ the impression management tactic of finding an exit strategy. In organizations, an exit strategy may be as explicit as stepping down from an executive position, quitting a high-stakes competition, exiting a negotiation, or terminating a project. Exit strategies, however, may also take subtler forms, such as switching tasks, reallocating attention, or shifting effort across projects. We predict that, when an initiative (e.g., a task, project, contest, negotiation)

falters, individuals facing high expectations will be more likely to employ an exit strategy than individuals facing low expectations, for three reasons.

First, individuals who face relatively high external expectations may be more likely to quit following poor initial performance to alleviate the distress caused by performing below expectations. For example, switching to a different task allows individuals with high external expectations to avoid direct exposure to the environment that has caused them embarrassment, and it can afford individuals the opportunity to validate themselves in another arena. This prediction is consistent with past research showing that embarrassment can motivate individuals to avoid the conditions that elicit embarrassment (e.g., by avoiding being observed by others or increasing psychological distance from others; Dong, Huang, & Wyer, 2013; Keltner & Buswell, 1997).

Second, individuals who face relatively high external expectations may be more likely to quit following poor initial performance to avoid the experience and attribution of a complete failure. Though quitting eliminates the opportunity to overcome an initial failure, it preserves the counterfactual that the quitter might have succeeded had she persisted. That is, quitting represents an extreme form of self-handicapping—an impression management tactic whereby individuals engage in activities that undermine their ability to succeed and thereby create an external attribution for poor performance (Berglas & Jones, 1978; Greenberg, Pyszczynski, & Paisley, 1984). Consistent with this prediction, past research has shown that, when people anticipate a potential failure, they sometimes engage in self-handicapping preemptively as a precaution (Greenberg et al., 1984). Also, in line with our theorizing, Kolditz and Arkin (1982) found that people are more likely to self-handicap when they anticipate performing for an audience with high expectations.

Third, in practice, the decision to reduce persistence (e.g., by quitting) after initial poor performance is often accompanied by an excuse. Executives who retire may cite the pressing need to spend time with family, politicians may cite health concerns, and athletes may cite an injury. By construction, excuses attribute poor performance to something other than a lack of skill and create an effective impression management strategy in that observers cannot easily determine the cause of the poor performance. Legitimate excuses preserve the impression that someone with poor initial performance would have performed better had it not been for extenuating circumstances (e.g., family obligations, health issues). By reducing persistence and offering an excuse, individuals make

the link between their poor performance and their true ability less clear (Berglas & Jones, 1978), thus diminishing the harm to their image caused by poor initial performance.

HYPOTHESES AND THEORETICAL CONTRIBUTIONS

Although past research has demonstrated that high external performance expectations can benefit individuals in many ways (e.g., by increasing self-expectations, motivation, and performance [Eden, 1984, 1990; McNatt, 2000]), scholars have failed to investigate the effect that high external performance expectations have on those who exhibit poor initial performance. We postulate that high expectations can burden individuals who encounter early setbacks, because poor initial performance creates impression management concerns and can trigger embarrassment. Following poor initial performance, we expect individuals with high expectations to be more likely to reduce persistence than individuals with low expectations because quitting can diminish embarrassment, prevent the experience of and attributions associated with a complete failure, and, if accompanied by a plausible excuse (e.g., the need to spend time with family, or an injury), offer an explanation for initial poor performance. Formally, we hypothesize:

Hypothesis 1. Individuals facing higher external performance expectations will be less persistent than those facing lower external performance expectations following poor initial performance.

Hypothesis 2. Individuals facing higher external expectations will experience more embarrassment than those facing lower external expectations following poor initial performance.

Hypothesis 3. Embarrassment will mediate the relationship between external expectations and persistence decisions following poor initial performance.

Our research makes several important theoretical contributions. First, we provide insight into how people make decisions about whether to persist or quit by identifying an often-overlooked motive that undermines persistence: impression management. That is, in addition to pursuing performance goals, individuals simultaneously pursue image goals that can predictably and systematically harm persistence following poor initial performance. Our findings challenge existing research that has presumed people decide to persist or quit based solely on their chances of success (e.g., Bandura, 1997; Berger & Pope, 2011; Carver, Blaney, & Scheier, 1979).

Second, our research contributes to prior work that has focused on the benefits of high performance expectations (e.g., Davidson & Eden, 2000; Eden, 1990; McNatt, 2000) and the disadvantages of low performance expectations (Babad et al., 1982). Surprisingly little research has considered the potential for harmful effects of high external expectations. A few studies, however, have shown that salient external expectations can impair performance by creating performance pressure and reducing the ability to concentrate (Baumeister, Hamilton, & Tice, 1985; Baumeister & Steinhilber, 1984; Cheryan & Bodenhausen, 2000). Our research documents a novel burden of high performance expectations and addresses the call for research to explore the downsides of positive expectations (Jensen & Kim, 2015; Kovács & Sharkey, 2014). Taken together, our findings advance our understanding of persistence as well as the psychological processes shaped by performance expectations.

OVERVIEW OF STUDIES

We report results from a field study (Study 1) and a laboratory experiment (Study 2) that investigate how external performance expectations influence individuals' persistence decisions and their experience of embarrassment, an emotion triggered by impression management concerns. We begin by examining a high-stakes, natural setting in which external expectations and the quality of initial performance vary substantially: professional men's tennis. We analyzed data on all available men's professional tennis matches played between 1973 and 2011. We used a regression discontinuity (RD) design to causally test whether being expected to win (vs. being expected to lose) increases the likelihood of quitting after falling behind in a match (Hypothesis 1). In Study 2, we experimentally manipulated external expectations and measured individuals' decisions to persist in a trivia challenge following poor initial performance. Through Study 2, we conceptually replicate our key findings from Study 1 in a different environment to establish the generalizability of the phenomenon and, importantly, we show that embarrassment mediates the relationship between external expectations and persistence in the face of adversity (Hypotheses 2 and 3).

STUDY 1: A FIELD STUDY OF PERFORMANCE EXPECTATIONS AND PERSISTENCE

In Study 1, we analyze behavior in a highly competitive setting: professional men's tennis. In this setting, individual competitors face clear external

performance expectations because their relative world rankings indicate whom the public expects to win any given match. Individual competitors make persistence decisions when they choose whether or not to quit during a match. Both world rankings and quitting decisions are readily observable and highly salient to competitors and observers.

For professional tennis players, quitting and losing yield identical world ranking and tournament outcomes (ATP Official Rulebook, 2018). However, compared to experiencing a loss, quitting has potentially different impression management consequences. When professional tennis players quit, they need to state a reason for quitting. The only officially accepted reason to quit a tournament is due to injury. That is, when players quit, they offer an excuse (in this case, poor health), which also provides an excuse for any poor prior performance.

Dataset

We compiled our data from the online data archive maintained by the Association of Tennis Professionals (ATP) World Tour (www.atpworldtour.com). Our dataset includes 328,515 men's professional singles tennis matches played between 1973 and 2011 that had information about both players' world rankings. Men's professional tennis matches nearly all consist of a best of two-out-of-three sets competition. For each match, we know the outcome of the match as well as the number of games won and lost by each player in each set, but we do not know the points that players won or lost in the games that made up sets, or the order in which games were won or lost during a set. Our dataset has information about whether either player quit mid-match, which we used to measure a player's persistence, as well as the match score at the time when a player quit. Our dataset also includes 1,849 matches that did not actually start or ended before any score was recorded due to a player withdrawing prior to a match or for other rare reasons (e.g., suspensions).

Analysis Strategy

We analyzed these data to test our prediction that, holding all else equal (e.g., player quality/skill, age, etc.), higher external performance expectations decrease persistence (i.e., increase the likelihood of quitting mid-match) after poor initial performance (Hypothesis 1). To assess whether or not a player exhibited poor performance early on in a match and thus faced a probable loss, we used an objective measure of initial performance. In tennis, players who

lose the first set of a match have a low probability of winning the match, on average. Across all matches in our dataset, players who lost the first set won the match only 18% of the time. Thus, we used defeat in the first set to indicate poor initial performance, and we excluded 4,218 matches from our analyses in which players did not finish the first set.

To assess external expectations faced by tennis players, we took advantage of the fact that observers form clear expectations of players' performance based on players' ATP world rankings. In a given match, the player with a superior ATP ranking is expected to win (i.e., faces higher performance expectations) and is called the "favorite," whereas the player with an inferior ATP ranking is expected to lose (i.e., faces lower performance expectations) and is called the "underdog." Confirming the validity of these rankings and expectations, we observed that favorites are more likely to win matches than underdogs (overall, favorites win 61.5% of matches).

In this study, we examined whether a tennis player who loses the first set in a match is more likely to quit when he is classified as a favorite than when he is classified as an underdog (Hypothesis 1). Identifying the impact of being a favorite rather than an underdog on a player's likelihood of quitting is not a simple task. This is because any correlation between a player's favorite (vs. underdog) status and his likelihood of quitting could be explained by omitted variables (e.g., players' relative skill levels) or self-selection issues (e.g., a player with an injury is more likely to show up to a match if his ranking is better than his opponent's). Ideally, we would test our hypothesis under conditions in which the player who is the favorite to win a match was randomly assigned; of course, this never happens. To approximate random assignment to favorite status, we used a quasi-experimental sharp RD design. An RD design involves assigning individual observations to a treatment or control group based on a continuous assignment variable (Imbens & Lemieux, 2008). Those observations above a discrete threshold of interest on the assignment variable are assigned to the treatment group, while others are assigned to the control group. The RD design examines an arbitrary threshold of theoretical interest to explore whether or not a stark discontinuity in outcomes (that otherwise change along a smooth continuum) emerges at the threshold. Because of its reliance on randomness, the RD design allows researchers to draw causal inferences about interventions and rule out self-selection or omitted variables as an alternative explanation for treatment effects (Imbens & Lemieux, 2008).

Our study's RD design involved assigning players to either favorite or underdog status based on a continuous

assignment variable that reflected players' likelihood of winning a match and captured players' relative skills. The threshold we relied upon in our study was the discontinuity that separates slight underdogs (players with rankings slightly worse than their opponents') from slight favorites (players with rankings slightly better than their opponents'). The continuous assignment variable we relied on is a transformed *rank ratio* variable that we calculated for each player in each match, amounting to the target player's opponent's ranking divided by the target player's own ranking. We took the logarithm transformation of this *rank ratio* to derive a measure we referred to as a player's *log rank ratio*. For example, a player ranked 100 facing an opponent ranked 101 will have a *log rank ratio* of $0.0100 = \log\left(\frac{101}{100}\right)$, and his opponent will have a *log rank ratio* of $-0.0100 = \log\left(\frac{100}{101}\right)$. A *log rank ratio* greater than 0 indicated that the target player had a better (i.e., lower) ATP ranking than his opponent, had higher external expectations for winning, and was classified as the favorite; a *log rank ratio* of less than 0 indicated that the target player had a worse (i.e., higher) ATP ranking than his opponent, had lower external expectations for winning, and was classified as the underdog. Importantly, the relationship between *log rank ratio* and a player's likelihood of winning a match was positive and continuous at the rank equality threshold (see Figure 1, Panel A), which is the reason we constructed and relied upon this measure.¹

The RD design allows us to compare tennis players with very similar skill levels and examine the causal effect of being classified as a favorite (or underdog) on a player's likelihood of quitting a match after poor initial performance. Players very near the rank equality threshold, but just over the threshold, are practically identical in terms of skill level. For example, imagine a match with players ranked 100 and 101. In our design, we assigned the player ranked 100 "favorite" status and his opponent ranked 101 "underdog" status. This assignment to favorite or underdog status is effectively random, since small differences in rankings are not

reliable signals of relative skill. However, prior research has demonstrated that people interpret ranking differences like these as if they were meaningful (e.g., Isaac & Schindler, 2014; Luca, 2014; Luca & Smith, 2013; Pope, 2009). For example, the observers of competitions are often overly sensitive to small differences in rankings. They base their forecasts of outcomes on rankings even when (a) rankings provide no additional information beyond the underlying performance scores used to generate them (e.g., Luca, 2014; Pope, 2009) and (b) changes in rankings are not accompanied by changes in objective quality measures (e.g., Luca & Smith, 2013). In addition, people are highly sensitive to boundaries in rankings (Isaac & Schindler, 2014). Altogether, this past research suggests that observers of competitions view small ranking differences as meaningful signals of differences in quality (or winning likelihood), even when this is inaccurate. We therefore expect small differences in rankings to inform external expectations, and in turn significantly and discontinuously influence players' persistence decisions. If we find this pattern in our data, we can conclude that being assigned favorite or underdog status exerts a causal influence on persistence.

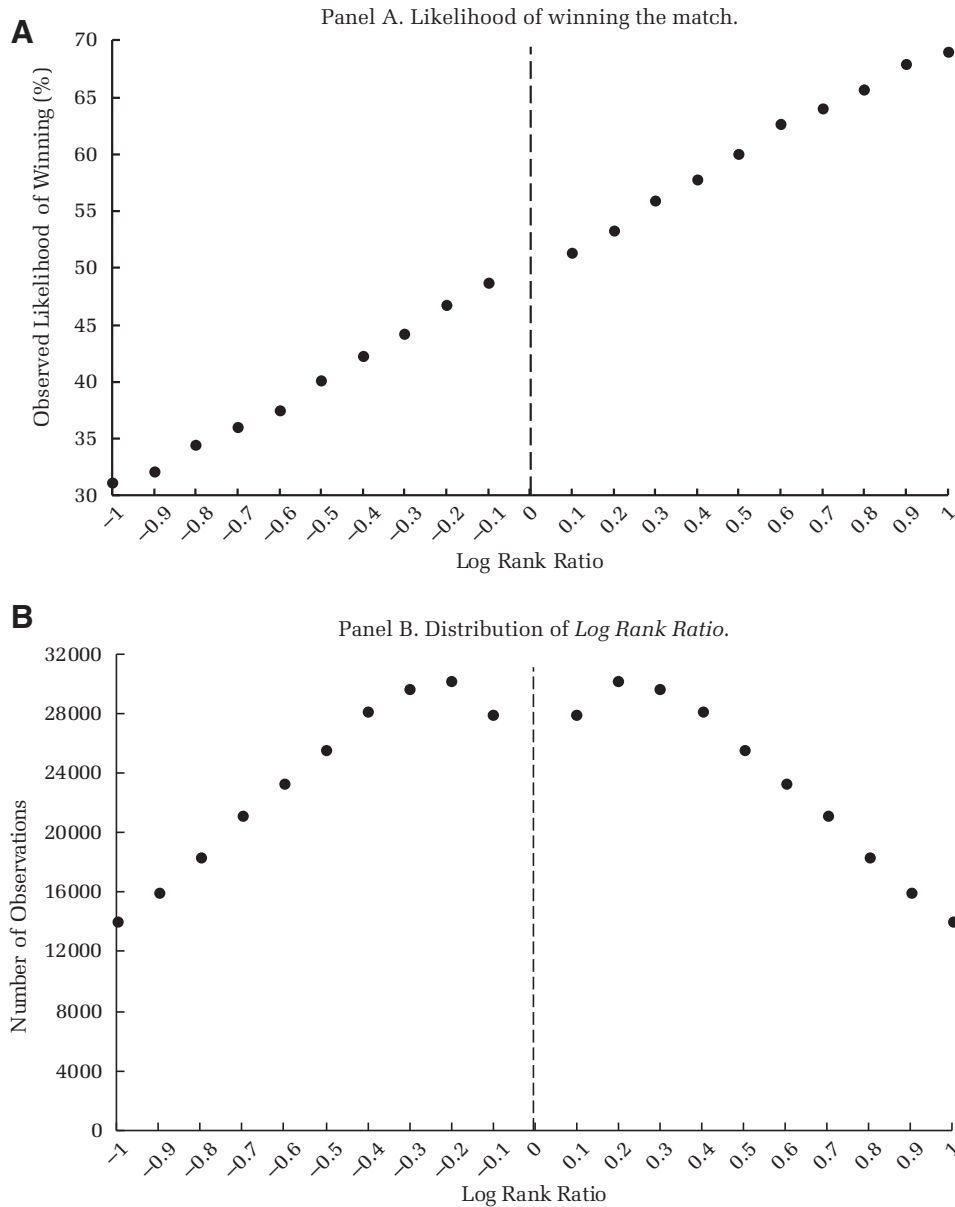
Variables

Our study's dependent variable, *quit*, is a binary variable that took on a value of 1 when a given player quit in a given match, and 0 otherwise. Quitting during a match is a stark measure of persistence. To capture whether external expectations were higher or lower for a player than for his opponent, we created an indicator variable, *favorite*, which equaled 1 if a player had a better ranking than his opponent and 0 otherwise. If the indicator variable *favorite* is a significant predictor of quitting, even after we account for players' relative skills using the continuous *log rank ratio* variable, we can conclude that being favored discontinuously affects a player's likelihood of quitting. Our dataset included 628 matches in which both players had the same recorded rank. In these rare cases, we were unable to label either player the favorite or underdog in a match. In our results, we excluded these matches, but our results were meaningfully unchanged if we treated both players in these matches as favorites or both as underdogs.

We included a number of important control variables in all of our analyses. First, following past research relying on RD designs, we controlled for a high-order polynomial of the assignment variable (e.g., a fourth-order polynomial in Flammer, 2015; a third-order polynomial in Pierce, Dahl, & Nielsen, 2013). Specifically, we included a fourth-degree

¹ Our regression analysis reported in Web Appendix B also confirms that there was not a discontinuity in the relationship between *log rank ratio* and winning likelihood around the rank equality threshold. This is important, because a discontinuity in the relationship between *log rank ratio* and winning likelihood near the rank equality threshold could offer an alternative explanation for RD analyses. Thus, we have focused on *log rank ratio*. However, our results are robust to using alternative continuous assignment measures such as the difference between two players' rankings or the ratio of two players' rankings (Web Appendix C).

FIGURE 1
Plots of the Likelihood of Winning a Match (Panel A) and Number of Observations (Panel B) as a Function of the Target Player’s *Log Rank Ratio* Relative to his Opponent



Notes: We depict matches in which players’ *log rank ratio* was within one standard deviation of the favorite–underdog threshold. For each of the 20 bins of width 0.1 of the *log rank ratio* variable, the observed likelihood of winning is calculated as the average likelihood of winning across all observations in that bin (Panel A), and the number of observations refers to the number of matches that fell in that bin (Panel B).

polynomial of *log rank ratio* to carefully control for the continuous relationship between a player’s skills relative to his opponent’s and his likelihood of quitting (and note that our results were robust to instead including a first-, second- or third-order polynomial).

Also, we controlled for the exact score in the first set by including fixed effects for first-set score (e.g., 0–6,

1–6, or 2–6, etc.) in order to compare players with the same outcomes in the first set of a match. In addition, we controlled for each player’s overall quality by including the focal player’s rank and his opponent’s rank. In order to account for possible differences in behavior at different rank levels, we also included interactions between a player’s rank and his opponent’s

rank with each term of the fourth-order polynomial of *log rank ratio*. Further, we controlled for each tournament’s tour (Grand Slam, Masters, etc.) and total available prize money in the tournament, because different types of tournaments may induce different levels of player motivation. Similarly, we included controls for the tournament round (first, second, third, etc.) in which a match was played because later rounds in a tournament typically have more spectators, prize money, and media coverage than earlier rounds. Different tours are characterized by different tournament sizes, so the same round number corresponds to a different number of remaining competitors across tours. To account for this, we included complete controls for the interactions between tour and round. Also, we controlled for court surface, which can affect injury (Girard, Eicher, Fourchet, Micallef, & Millet, 2007) and quitting rates (Breznik & Batagelj, 2012). Finally, we controlled for a player’s age, as age may relate to injury-proneness. In order to account for a possible nonlinear relationship between age and injury-proneness, we also included a squared term of player’s age. Table 1 shows summary statistics for all matches in our sample that actually started. Web Appendix A provides additional information about control variables in our regressions. Notably, our results are robust to excluding all of the control variables detailed in this paragraph.

Regression Specifications

We used an ordinary least squares (OLS) model to predict quitting,² though our results were meaningfully unchanged when we instead relied on logistic regression models (see results in Web Appendix C). Given our primary interest in examining the effects of being a favorite on quitting rates among players who exhibited poor initial performance, we only included the player from each match who lost the first set. Formally, our model can be stated as follows, for a target player *i* in a given match *j*:

TABLE 1
Descriptive Statistics

Variable	Mean	SD
Quit	0.014	0.117
Favorite	0.50	0.50
First-Set Loser	0.50	0.50
Log Rank Ratio	0.00	1.15
Target Player’s Rank	389.65	382.35
Opponent’s Rank	389.65	382.35
Age	24.10	3.72
Prize Money (\$)	245,374.20	660,006.90
Year	1998.83	9.92
Round	1.92	1.13
Win the Match	0.50	0.50

Notes: Descriptive statistics are calculated for 657,030 player–match observations from 328,515 matches with the exception of first-set loser for which description statistics are calculated for 648,594 player–match observations from 324,297 matches in which players completed the first set. Descriptive statistics summarizing categorical control variables (tour types, surface types, and first-set scores) are detailed in Web Appendix A.

$$\begin{aligned}
 quit_{ij} = & \alpha + \beta_1 * favorite_{ij} + \beta_2 * \log(rank\ ratio_{ij}) \\
 & + \beta_3 * \log(rank\ ratio_{ij})^2 + \beta_4 * \log(rank\ ratio_{ij})^3 \\
 & + \beta_5 * \log(rank\ ratio_{ij})^4 + \beta_6 * rank_{ij} \\
 & + \beta_7 * \log(rank\ ratio_{ij}) * rank_{ij} \\
 & + \beta_8 * \log(rank\ ratio_{ij})^2 * rank_{ij} \\
 & + \beta_9 * \log(rank\ ratio_{ij})^3 * rank_{ij} \\
 & + \beta_{10} * \log(rank\ ratio_{ij})^4 * rank_{ij} \\
 & + \beta_{11} * opponent\ rank_{ij} \\
 & + \beta_{12} * \log(rank\ ratio_{ij}) * opponent\ rank_{ij} \\
 & + \beta_{13} * \log(rank\ ratio_{ij})^2 * opponent\ rank_{ij} \\
 & + \beta_{14} * \log(rank\ ratio_{ij})^3 * opponent\ rank_{ij} \\
 & + \beta_{15} * \log(rank\ ratio_{ij})^4 * opponent\ rank_{ij} \\
 & + \beta_{16} * age_{ij} + \beta_{17} * age_{ij}^2 + \beta_{18} * prize\ money_{ij} \\
 & + \delta * X_{ij} + \theta * Z_j + \epsilon_{ij}
 \end{aligned}
 \tag{1}$$

where *quit_{ij}* took a value of 1 when target player *i* quits in match *j*, and 0 otherwise. Our primary predictor variable was the indicator for being perceived as a favorite (vs. underdog). We expect the coefficient on this variable to be significant and positive among players who lost the first set. *X_{ij}* is a vector of control variables representing the first-set score from the target player’s perspective. For example, if the target player won six games and his opponent won three games in the first set, the first-set score would be

² We report the result from OLS regression models in our primary analyses rather than logistic regressions for several reasons. Not only are results from OLS regressions easier to interpret, but we include a large number of fixed effects in our models, and logistic regression models typically produce inconsistent estimates when fixed effects are included unless data characteristics meet a stringent set of assumptions (for details about the “incidental parameter problem,” see Wooldridge, 2010). However, as shown in Web Appendix C, our results are nearly identical when we instead rely on logistic regressions.

coded as 6–3 for the target player. \mathbf{Z}_j is a vector of control variables representing tournament round, tour, round–tour interaction, court surface, and year corresponding to the match j . Finally, ε_{ij} is the error term. To account for potential within-player dependence across observations, we clustered standard errors at the player level.

Results

The average rate of quitting mid-match is 1.39% in our dataset, which means quitting occurred in 9,132 matches. Figure 2 depicts quitting rates for matches in which players' *log rank ratio* was within one standard deviation of the rank equality threshold (i.e., *log rank ratio* between -1 and 1). We group observations into bins based on players' *log rank ratio* at intervals of 0.1 (i.e., 10% of the standard deviation of *log rank ratio*). Each dot represents the average rate at which players in the designated bin quit in the middle of a match. Figure 2 (Panel A) illustrates that, among players who exhibited poor initial performance and lost the first set of a match, there is a significant discontinuity in a player's likelihood of quitting at the threshold separating favorites from underdogs. This pattern is consistent with Hypothesis 1. Slight underdogs whose *log rank ratio* falls between -0.10 and 0 (i.e., whose rank is approximately 90–100% of their opponents' rank; $n = 14,072$) retire 1.72% of the time, but slight favorites whose *log rank ratio* is between 0 and 0.10 (i.e., whose rank is approximately 100–110% of their opponents' rank; $n = 13,575$) retire at a significantly higher rate: 2.04% of the time, $\chi^2(1) = 3.86, p < .05$.³ The increase in the likelihood of quitting is 18.60% when a player moves to the favorite side of this favorite–underdog threshold, and this is notably larger than the increase we observe at other arbitrary thresholds in Figure 2.

We next report a series of regression models that take into account relevant control variables to test Hypothesis 1. Model 1 in Table 2 follows regression

specification (1), which focuses on players who exhibited weak initial performance and lost the first set. The positive and significant coefficient estimate on *favorite* ($\beta = 0.0037, p < .001$) indicates that players who lost the first set have markedly different patterns of quitting behavior right at the threshold of equal rank, as predicted by Hypothesis 1. Specifically, being a slight favorite is associated with a 0.37-percentage-point increase in the probability that a player will quit mid-match (or an increase of 26.62% relative to the average quitting rate observed in our data).

Testing the validity of the RD design. Thus far, we have relied on an RD design to show that being the favorite (rather than the underdog) in a competition significantly increases a player's likelihood of quitting when he exhibits poor initial performance. To claim that we have *causal* evidence of this fact, we need to ensure that we satisfy our RD design's two critical assumptions (Imbens & Lemieux, 2008): (1) around the equal rank threshold, being assigned to favorite versus underdog status is essentially random and (2) there is not a discontinuity in the distribution of *log rank ratio* at the rank equality threshold. We conducted standard tests to examine these assumptions, and present detailed results in Web Appendix B.

With respect to the first assumption, we confirmed that the characteristics of players immediately above and below the equal rank threshold are very similar. Importantly, we found that the likelihood of winning the first set in a match or of winning an entire match does not significantly differ between players who are immediately above and those who are immediately below the equal rank threshold, confirming that slight favorites and slight underdogs have comparable levels of skill.

With respect to the second assumption, we grouped player–match observations into bins based on their *log rank ratio* and plotted the number of observations in each bin (Imbens & Lemieux, 2008). We showed that the distribution of *log rank ratio* is symmetrical around the rank equality threshold (Figure 1, Panel B), which must be true by definition of the way underdogs and favorites are defined. The ATP pairs players as randomly as possible, except that competition between players who are both seeded (and often have similar rankings) in a tournament is avoided by the ATP in early rounds of tournaments. This reduces the number of observations as *log rank ratio* becomes very close to the rank equality threshold, which

³ When we instead examine slight underdogs whose *log rank ratio* falls between -0.15 and 0 (i.e., whose rank is approximately 85–100% of their opponents' rank; $n = 21,864$) and compare them to slight favorites whose *log rank ratio* is between 0 and 0.15 (i.e., whose rank is approximately 100–115% of their opponents' rank; $n = 20,636$), we again see that favorites retire at a significantly higher rate (2.09%) than underdogs (1.62%), $\chi^2(1) = 12.91, p < .001$.

explains why the distribution has a dip right around the rank equality threshold.⁴

The concern about a possible discontinuity in the distribution of *log rank ratio* at the rank equality threshold could also arise if players self-select from one side of the favorite–underdog threshold to the other. However, players cannot choose their rankings, their opponents, or the structure of tournament draws, all of which are decided by ATP officials. The only way self-selection could possibly occur is if, after being assigned an opponent but before beginning a match, some players choose to be “no shows” and withdraw (hereafter, *pre-match withdrawals*). The average rate of pre-match withdrawals was 0.19% in our dataset. It is, in principle, possible that a discontinuity in pre-match withdrawals between underdogs and favorites could explain the discontinuity we detected in mid-match quitting. We therefore conducted regression analyses to predict a player’s choice to withdraw prior to a match as the binary dependent variable, and we confirmed that there was not a discontinuity in pre-match withdrawals between slight favorites and slight underdogs (see Web Appendix B).

Ruling out discontinuities at other thresholds as an explanation for quitting behavior. We next considered whether or not quitting was so unpredictable that statistical tests spuriously revealed discontinuities in quitting rates not only at the favorite–

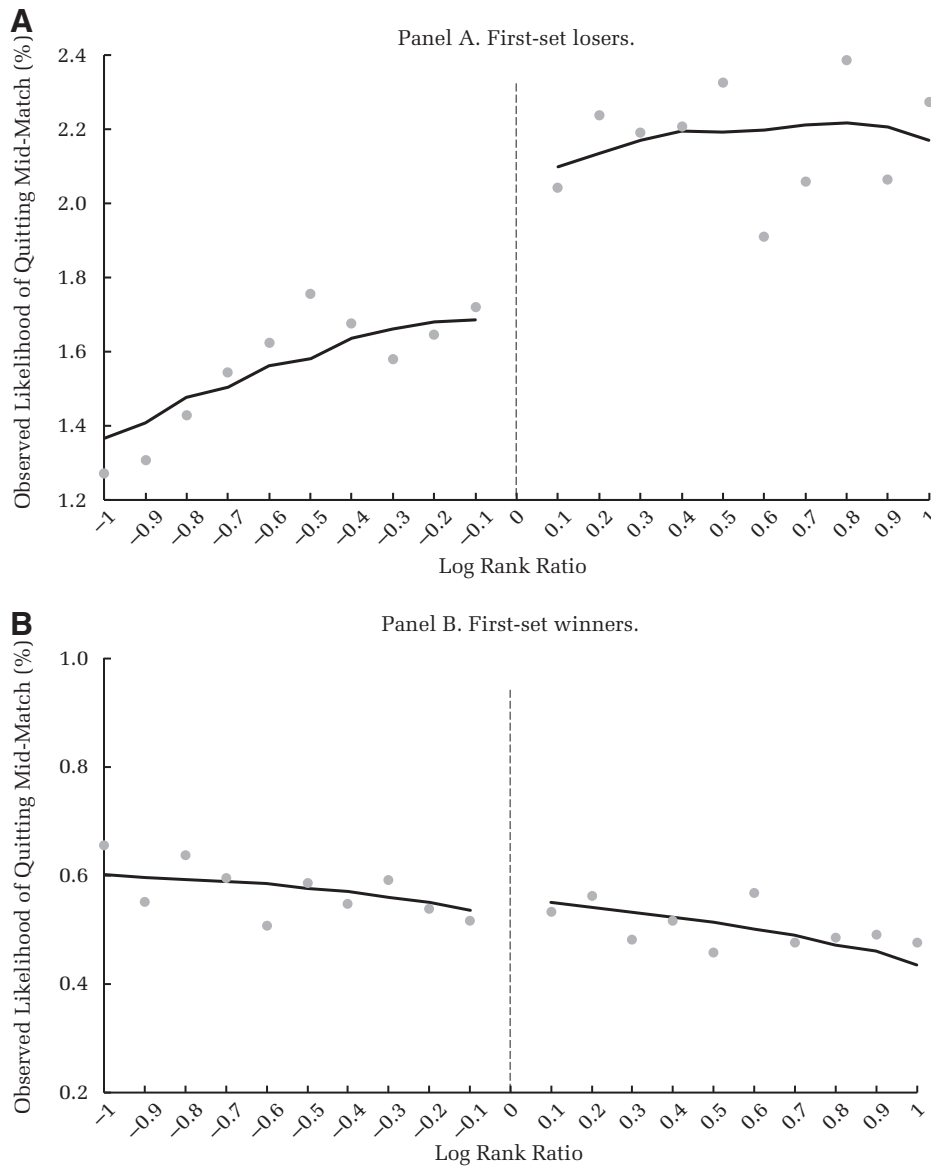
underdog threshold, but at many other points along the *log rank ratio* continuum. The importance and interpretability of the discontinuity we detected in quitting rates at the favorite–underdog threshold would be diminished if there were discontinuities in quitting at other, less meaningful thresholds. This is because discontinuities at other locations would suggest that our key findings might reflect an artifact of data irregularities or of the statistical tests we used rather than clear support for Hypothesis 1.

To rule out this possibility, we followed the procedures described by Pierce et al. (2013) and reran the regression specified in Model 1 of Table 2 21 times with slight alterations. Specifically, we replaced the favorite indicator with an indicator for whether a player’s *log rank ratio* was greater than a new threshold of X , where X was increased in each iteration by 0.1 from a minimum of -1.0 to a maximum of 1.0 . For $X = 0$, this regression is the same regression shown in Model 1 of Table 2, which tests whether being assigned favorite status significantly increases the likelihood of quitting after a player loses the first set. Figure 3 presents the estimated coefficients on the modified favorite indicator for each of these 21 regressions together with 95% confidence intervals. The largest and most significant coefficient on our favorite indicator arises at the threshold of interest—the salient tied-rank threshold (*log rank ratio* = 0). When the threshold value is adjusted upward or downward one “step” from this salient threshold and set to -0.1 or 0.1 , we also observe a statistically significant discontinuity, which is very likely driven by the close proximity between these “placebo” threshold values and the rank equality threshold. No other placebo threshold values generate significant discontinuities. Thus, these placebo tests suggest that the psychologically meaningful threshold of rank equality has a unique effect on players’ quitting behavior.

Ruling out discontinuities following strong initial performance. Our theory predicts that, after exhibiting poor initial performance, individuals facing high external expectations become less persistent than individuals facing low external expectations due to impression management concerns. However, we do not expect this difference to emerge when initial performance is strong: strong initial performance is unlikely to elicit impression management concerns regardless of individuals’ prior expectations. Thus, we conducted additional analyses to investigate whether our finding that

⁴ When we focus on a narrower window around the rank equality threshold (where the *log rank ratio* falls between -0.2 and 0.2) and use a smaller bin size of 0.01 (Web Appendix B), we observe a smooth distribution, except that the number of observations in a bin drops when the absolute value of *log rank ratio* changes from being between 0.01 and 0.02 to being between 0 and 0.01 . Our results reported in Table 2 remain robust if we remove matches in which competitors’ *log rank ratio* falls between -0.01 and 0.01 ($N_{\text{match}} = 2,593$). Also, looking at raw data, we see that slight favorites whose *log rank ratio* is between 0.01 and 0.10 ($n = 12,658$) retire at a significantly higher rate (2.01%) than slight underdogs whose *log rank ratio* falls between -0.10 and -0.01 ($n = 13,104$; quitting rate = 1.67%), $\chi^2(1) = 4.02, p < .05$. Furthermore, when we focus on an even smaller region around the rank equality threshold from -0.02 to 0.02 and use a smaller bin size of 0.001 (Web Appendix B), we see a continuous distribution of *log rank ratio* around the threshold (instead of a sharp drop at the threshold). In addition, if we use the difference between two players’ rankings as the assignment variable, the number of observations in each bin has a smooth distribution, and, importantly, the quitting results hold. See Web Appendix C for details.

FIGURE 2
The Relationship between a Target Player’s Log Rank Ratio Relative to his Opponent and the Target Player’s Likelihood of Quitting Mid-Match as a Function of First Set Outcomes



Notes: We depict matches in which players’ *log rank ratio* was within one standard deviation of the favorite–underdog threshold. Observations are grouped into bins of width 0.1 based on the *log rank ratio* variable. Panel A shows outcomes for first-set losers; Panel B, for first-set winners. Each dot represents the average rate at which players in the designated bin quit in the middle of a match. Solid lines depict the fitted probabilities calculated by taking the mean of all predicted values from the tested models (Panel A from Model 1 in Table 2, and Panel B from Model 2 in Table 2) for observations in a given bin.

favorites quit at a higher rate than underdogs is unique to situations in which initial performance is poor.

First, we relied on regression specification (1) to predict quitting among players who *won* the first set. Model 2 in Table 2 suggests that favorites are no more likely to quit than underdogs if they

exhibited strong initial performance and won the first set ($\beta = 0.0002, p = .72$). This is consistent with what we observe through visual inspection of the raw data: Figure 2 (Panel B) shows that, within the narrow window around the favorite–underdog threshold (i.e., *log rank ratio* between -0.1 and 0.1), slight underdogs and slight favorites exhibit

TABLE 2
OLS RD Models Predicting Quitting Mid-Match as a Function of Whether a Player was the Favorite to Win a Match

Outcome Variable	Quitting Mid-Match		
	Model 1	Model 2	Model 3
Favorite	0.0037*** (0.0010)	0.0002 (0.0005)	-0.0005 (0.0009)
First-Set Loser × Favorite			0.0053*** (0.0016)
First-Set Loser			0.0296*** (0.0033)
Log(Rank Ratio)	0.0006 (0.0006)	-0.0003 (0.0004)	-0.0010 (0.0006)
Log(Rank Ratio) ²	-0.0003 (0.0002)	0.0001 (0.0001)	0.0007*** (0.0002)
Log(Rank Ratio) ³	4.48E-06 (4.32E-05)	-7.46E-05 (4.82E-05)	2.70E-05 (6.05E-05)
Log(Rank Ratio) ⁴	-2.37E-06 (1.14E-05)	7.88E-06 (1.14E-05)	-2.60E-05 (1.52E-05)
First-Set Loser × Log(Rank Ratio)			0.0024* (0.0010)
First-Set Loser × Log(Rank Ratio) ²			-0.0011*** (0.0003)
First-Set Loser × Log(Rank Ratio) ³			-8.68E-05 (7.56E-05)
First-Set Loser × Log(Rank Ratio) ⁴			2.83E-05 (1.73E-05)
Additional control variables	Yes	Yes	Yes
Sample inclusion criterion	Players who lost the first set in a match are included	Players who won the first set in a match are included	One player is randomly selected from each match
Observations	323,683	323,683	310,844–340,739 across 100,000 samples
R ²	0.0057	0.0017	0.0066–0.0080 across 100,000 samples

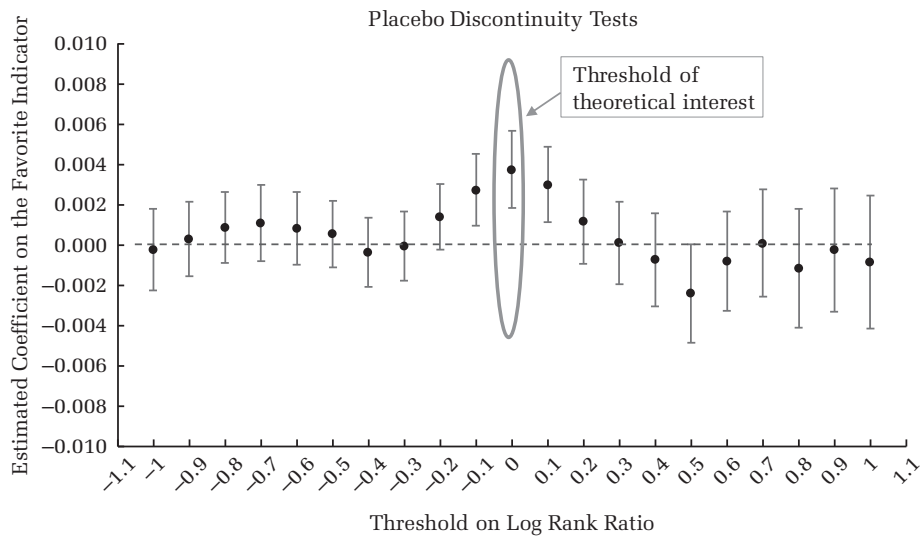
Notes: Model 1 includes players who lost the first set in a match. It shows that favorites who lost the first set are more likely to quit than underdogs who lost the first set. Model 2 includes players who won the first set in a match. It shows that favorites who won the first set are no more likely to quit than underdogs who won the first set. Model 3 presents analyses that randomly select one player from each match without conditioning target player selection on first set performance. It shows that the discontinuous jump in quitting rates at the favorite–underdog threshold is significantly stronger for players who lose the first set than for players who win the first set. Standard errors are clustered at the player level in Models 1 and 2. For Model 3, we report the mean and standard deviations of the coefficients from 100,000 bootstrapped samples, and calculate *p* values based on bootstrapped standard errors. Additional control variables include player’s rank and its interactions with log rank ratio polynomial, opponent’s rank and its interactions with log rank ratio polynomial, player’s age and its squared term, prize money, match surface fixed effects, tour fixed effects, tournament round fixed effects, tour × round fixed effect interactions, first set score fixed effects, and year fixed effects. *, **, *** denotes significance at the 5%, 1%, and 0.1% levels, respectively.

virtually identical quitting rates after winning the first set (0.53% for slight underdogs and 0.52% for slight favorites, *p* = .84).

Next, we examined whether the effect of being a favorite on quitting significantly differs between players who had poor initial performance (those who lost the first set) and players who had strong initial performance (those who won the first set). This analysis required us to simultaneously examine the decisions made by two players per match. By construction, one player’s choice to quit precludes his opponent from quitting. To deal with the dyadic,

non-independent structure of our data as well as the non-independence across observations within each player, we combined a resampling methodology with block bootstrapping (Cameron, Gelbach, & Miller, 2008). Specifically, we created 100,000 bootstrap resamples in two steps. First, we randomly selected one observation (representing a player in a specific match) from each match to create a sample that was half of the size of our original dataset. We repeated this step 100 times. Then, within each sample constructed in the first step, we resampled the clusters of players with

FIGURE 3
Estimated Change in Quitting at Different (Placebo) RD Thresholds, as Well as the Threshold of Theoretical Interest



Notes: A fourth-order RD model (similar to Model 1 in Table 2) was run for each threshold value of *log rank ratio* between -1.0 and 1.0 at 0.1 intervals. Each dot represents the coefficient on the favorite indicator for a given threshold. Error bars represent 95% confidence intervals.

replacements 1,000 times and included all observations of players that were randomly selected. On each of the 100,000 (i.e., $100 \times 1,000$) bootstrap resamples, we ran an OLS regression to predict quitting using the independent variables in regression specification (1) along with an indicator for the loser of the first set in a given match (*first-set loser*) and its interaction with the *favorite* indicator. We also included the interaction between an indicator for being the *first-set loser* with each term in the fourth-degree polynomial of *log rank ratio* (four terms total) to ensure that the interaction between *first-set loser* and the *favorite* indicator was not picking up a continuous effect of *log rank ratio* on quitting but instead was reflecting a discontinuous difference between favorites and underdogs. For each bootstrap resample, we stored the resulting regression coefficients. We report the means and standard deviations of the coefficient estimates across 100,000 bootstrap resamples in Model 3 in Table 2. The primary predictor of interest is the interaction term between an indicator for losing the first set and the favorite indicator. The positive and significant coefficient on the interaction term ($\beta = 0.0053$, $p < .001$) indicates that the discontinuous jump in quitting rates at the favorite–underdog threshold is significantly stronger for favorites who lose the first set than for

favorites who win the first set, consistent with our theorizing.⁵

Robustness tests. We conducted numerous analyses to test the robustness of our findings pertaining to Hypothesis 1, which we present in Web Appendix C. To summarize these tests, our results remain robust when (1) we rely on logistic regressions instead of OLS regressions; (2) we adjust the operationalization of predictor variables, including (a) controlling for a player’s age as a linear term without the squared term, (b) treating both players in a match as favorites if they have identical rankings (in cases in which players have identical rankings, the average quitting rate among first-set losers is 1.95%), or (c) treating both

⁵ Summing the coefficient on the *favorite* indicator and the coefficient on the interaction term reveals that the total effect of being a favorite on quitting likelihood for players who lost the first set is statistically significant (the size of the total effect is 0.0049, with a bootstrapped *SE* of 0.0014, $p < .001$). This estimated effect size, within the margin of error, is similar to the effect estimated in Model 1 (i.e., 0.0037; $SE = 0.0010$) where only players who lost the first set were included. In addition, the estimated effect of being a favorite for players who won the first set was similar in Model 2 (coefficient on *favorite* = 0.0002, $SE = 0.0005$) and Model 3 (coefficient on *favorite* = -0.0005 , bootstrapped $SE = 0.0009$), within the margin of error.

players in a match as underdogs if they have identical rankings; (3) we remove matches with extremely high or low *log rank ratios* by dropping all observations with a *log rank ratio* above the 97.5th or below the 2.5th percentile; (4) we remove matches with a *log rank ratio* between -0.01 and 0.01 ; and (5) we use differences in players' ranks or players' raw rank ratio instead of their *log rank ratio* to capture relative skills.

It is common in RD analysis to control for high-order polynomials of the continuous assignment variable (e.g., Flammer, 2015; Lee & Lemieux, 2010; Pierce et al., 2013). However, Gelman and Imbens (2014) presented evidence that estimators based on high-order polynomial models can be misleading in some situations. We conducted several robustness tests to address this concern. First, our results remain meaningfully unchanged if we control for the second-order or third-order polynomial of *log rank ratio* rather than the fourth-order polynomial. In addition, our results are robust if we remove high-order polynomials altogether and only control for a linear term of *log rank ratio*. Further, our results are robust to applying local linear models to observations that are close to the rank equality threshold, as recommended by Gelman and Imbens (2014). Specifically, in a robustness test, we discarded observations with a *log rank ratio* value that was greater than an optimal bandwidth away from our threshold of 0 in order to avoid concerns about a non-linear relationship between the *log rank ratio* variable and the decision to quit (Angrist & Pischke, 2009; Imbens & Lemieux, 2008). We then estimated a linear function on remaining observations with *log rank ratio* values near 0 (Hahn, Todd, & van der Klaauw, 2001). We followed Imbens and Kalyanaraman (2009, 2012) to derive the optimal bandwidth around the rank equality threshold for this test. Hypothesis 1 remains supported when we switch to this alternative empirical approach. In Web Appendix D, we provide details about this analytical approach and the regression results it produces.

Discussion

In this field study, we analyzed the behavior of thousands of men's professional tennis players and identified a favorite-underdog discontinuity in their persistence decisions. Compared to underdogs who exhibit poor initial performance (falling behind by a set), favorites who exhibit poor initial performance are more likely to quit.

We found that being a slight favorite (vs. a slight underdog) increased the likelihood of quitting among first-set losers by 26.62% relative to the average quitting rate observed in our data. This detected effect was robust to numerous alternative specifications and robustness tests. Importantly, our RD design did *not* compare average favorites with average underdogs, and we did not rely on all of the 323,683 observations in which players lost the first set to identify the average effects of favorite status on quitting following a setback. In fact, when we zoomed in on matches involving similarly ranked players and considered the 27,647 observations in which players' *log rank ratios* fell between -0.10 and 0.10 ,⁶ we found that favorites quit significantly more often than underdogs. Further, in RD analyses including all 323,683 observations in which the target player lost the first set, we also identified our effects by comparing favorites and underdogs who were just above or below the favorite-underdog threshold thanks to the inclusion of a high-order polynomial of the *log rank ratio* variable in our regression models. Our regression approach including all observations and a large set of control variables allowed us to rule out the possibility that underlying differences between underdogs and favorites (e.g., in wealth, number of fans, skills, etc.) could account for our findings. We can also rule out selection effects for our findings since tennis players cannot directly influence the opponent they face, and we observed no differences in pre-match withdrawals between underdogs and favorites. Overall, our field study leverages an RD design to make a causal inference about the effect of being favored on quitting, and supports Hypothesis 1.

⁶ We balance two considerations when choosing bandwidths for our analyses. Choosing too wide of a bandwidth raises the concern that we might be comparing dissimilar players. Choosing too narrow of a bandwidth, however, raises the concern that we include too few instances of quitting and lack statistical power to detect a difference between slight favorites and slight underdogs. We conducted our analysis on players whose *log rank ratios* fell between -0.10 and 0.10 to balance these concerns. The extremes of the *log rank ratio* in this region (i.e., -0.10 and 0.10) correspond to a #20 player playing a #18 player or a #22 player, which we think represent skill levels that are quite comparable. If we use a smaller bin size (e.g., 0.01 ; see Figure B4 in Web Appendix B), the quitting results become noisier, because there are not a large enough number of observations extremely close to the rank equality threshold ($n = 1,885$ observations in which the target player lost the first set and his *log rank ratio* fell between -0.01 and 0.01) and incidences of quitting are relatively rare (the averaging quitting rate is 1.68% among first-set losers).

According to ATP rules, the only legitimate reason a player has for leaving a match early is injury. Thus, we know that each player who quit in the middle of a match claimed to be injured. In our dataset, slight favorites were discontinuously more likely to quit and claim an injury than slight underdogs after losing a first set, even though these competitors were virtually identical (e.g., in terms of their likelihood of getting injured and the seriousness of their injuries) beyond their designations as “favorite” and “underdog.” Thus, we have suggestive evidence that some of the injuries that favorites claimed were, at the very least, less severe than those claimed by underdogs, if not entirely fictitious. This is consistent with our proposed mechanism that people use quitting as an impression management strategy. However, there may be other mechanisms that contribute to our findings about tennis players’ quitting decisions. For example, favorites who are down in the first set may try harder to catch up, take on more on-court risks, and thus have a higher risk of true injuries, compared with underdogs who are down in the first set. In Web Appendix E, we report additional analyses showing that this alternative explanation is unlikely to account for the effect of performance expectations on quitting that we document, however. Another potential alternative explanation is that favorites might interpret the same injuries as discontinuously more painful and serious than underdogs after losing the first set. Favorites may even be more likely to interpret their initial loss as a signal that they are injured, compared to underdogs. Such differences in attributions may cause favorites to quit more frequently than underdogs after losing the first set. Our field study cannot disentangle our hypothesized impression management motive (i.e., players quitting and falsely claiming injuries to avoid further public embarrassment) from these attribution-based explanations. In our next study, we report the results of a laboratory experiment designed to test our hypotheses and replicate our findings in a different setting. With this laboratory study, we highlight the generalizability of our findings and directly examine the mediating role of embarrassment on persistence decisions.

STUDY 2: EXPECTATIONS, IMPRESSION MANAGEMENT CONCERNS, AND PERSISTENCE IN THE LABORATORY

In Study 2, we extended our investigation to a different setting outside of the context of sports: we conducted a laboratory study involving a trivia challenge. In this study, we manipulated participants’ external performance expectations and measured their feelings

of embarrassment and their persistence decisions. Study 2 serves three primary purposes. First, we replicate our finding in Study 1 showing that individuals who face high performance expectations are less persistent after poor initial performance (Hypothesis 1). Second, we offer insight into the underlying mechanism responsible for this effect; we demonstrate that, following poor initial performance, individuals who face high external expectations feel more embarrassed than individuals who face low external expectations (Hypothesis 2). Finally, we demonstrate that feelings of embarrassment mediate the influence of external expectations on persistence decisions (Hypothesis 3). Following best practices, we report how we determined our sample size, all data exclusions, all manipulations, and all measures.

Participants

We conducted this study at a Midwestern university in the United States. Participants received course credit for participating in the experiment and earned additional money depending on their performance. We recruited as many participants as we could across 29 experimental sessions, and the number of sessions was determined prior to data collection. Among the 305 participants who showed up for these 29 sessions, 304 participants (161 females, average age of 19 years) completed our study.⁷ The number of participants who attended each session ranged from six to 12.

Experimental Procedure

At the start of each experimental session, the experimenter asked participants to introduce themselves to one another by stating their names and academic majors (Chen et al., 2011; Lim, 2010). Then, the experimenter handed out and read aloud instructions about a trivia challenge. Participants learned that they would be randomly assigned to take part in a trivia challenge including questions of either “middle-school difficulty” or “expert difficulty.” Participants learned that the *middle-school difficulty* condition involved trivia questions to which most people are exposed by the end

⁷ Due to technical difficulties with one computer, one participant could not proceed after the 20th trivia question and did not answer any subsequent questions (including questions used to measure the underlying mechanism responsible for persistence). Our results regarding the relationship between performance expectations and persistence are robust if we include this participant and assign them either the lowest (i.e., 20) or the highest (i.e., 45) possible persistence score.

of middle school, and that good performers answer 70% of the questions correctly. Participants also learned that the *expert difficulty* condition involved trivia questions that most people cannot answer, and that good performers answer 20% of these questions correctly.⁸

By assigning participants to the middle-school difficulty or expert difficulty condition, we manipulated whether participants faced external expectations to perform at a relatively high or low level, respectively. Since all participants read the same instructions, external expectations for both groups were common knowledge. Unbeknownst to the participants, the questions were identical between conditions. Through pretests, we selected questions that most people would have been educated about by the end of middle school but found difficult (e.g., “Where does the presentation of the Nobel Peace Prize occur annually?”; see Web Appendix F).

We informed participants that they would have up to 30 seconds to answer each question. We also told participants that the first 20 trivia questions would cover a broad array of topics, and that, after the 20th question and following every question thereafter, they would be allowed to choose whether to continue answering trivia questions covering the same set of topics or to switch to trivia questions on new topics. We told participants that, regardless of whether or not they switched topics, the difficulty level of the questions would remain the same.

Next, we informed participants that their performance (i.e., the percentage of questions they answered correctly) would be ranked relative to the performance of other participants in the same trivia difficulty group and posted on the blackboard before the experimental session ended. This aspect of our design made participants’ performance public and created social performance pressure (because peers and experimenters observed their performance). Participants learned that, if they switched topics at any point, their performance would be calculated based only on the questions they answered after switching

topics. Thus, by deciding to quit answering questions on the original topics, participants could reset their performance and prevent past (potentially poor) performance from being posted publicly.

Finally, we informed participants about their incentives: participants who persisted in answering quiz questions on the original set of topics would earn \$0.15 for each correct answer during the trivia challenge; participants who switched topics at any point would earn \$0.15 for each correct answer submitted before they switched topics and \$0.10 for each correct answer they submitted after they switched. Thus, participants faced a monetary penalty for failing to persist. If they switched topics, they would sacrifice potential earnings (earning \$0.10 instead of \$0.15 per correct question), but reset their performance.

After the experimenter finished reading instructions to participants and answering their questions, she flipped a coin to determine each participant’s condition. Participants in the middle-school difficulty condition sat at computer terminals labeled “Trivia Challenge: Middle-School Difficulty”, while participants in the expert difficulty condition sat at computer terminals labeled “Trivia Challenge: Expert Difficulty.”

After each question was answered, we gave participants feedback so that they knew whether or not their answer was correct and what percentage of questions they had answered correctly up to that point in the study. After participants received feedback on the 20th question, they answered a series of questions that assessed how they felt about their performance and the results they had achieved from the challenge. We detail these questions in the Measures section, below.

Next, we presented participants with their first opportunity to quit answering questions on the original set of topics and switch to questions about a new set of topics (and earn a lower piece-rate payment per correct answer). We offered participants who continued with the original topic set the opportunity to switch after every question until the end of the challenge (which included 45 questions on the original topics). Notably, we did not inform participants how many questions were included in the challenge or what new topics they would encounter if they chose to switch topics.

After finishing the trivia challenge, participants answered questions about their age and gender and finally received their payment. Participants’ performance was ranked and posted on the blackboard at the end of the session.

⁸ Our decision to use 70% and 20% in our communications manipulating external expectations was informed by a pretest ($n = 213$). In the pretest, participants only read descriptions of the two conditions without information about the results of “good” performers. Participants assigned to the middle-school difficulty condition predicted that they would be able to answer 70% of the questions correctly on average, whereas participants actually answered 20% of questions correctly on average in both conditions.

Measures

Our dependent measure of persistence was the number of questions from the original topic set that each participant attempted to answer before electing to switch topics. This measure took on a value ranging from 20 (for participants who switched topics at the first opportunity) to 45 (for participants who never switched topics).

After the 20th trivia question, we asked participants to rate on a seven-point Likert scale the extent to which they felt embarrassed (1 = *not at all*; 7 = *very much*) using a four-item scale adapted from Mosher and White (1981) (“embarrassed,” “self-conscious,” “ashamed,” and “disgraced”). Though some prior scholars have distinguished shame from embarrassment, the four items were highly correlated in this study (Cronbach’s $\alpha = .89$). Thus, we averaged responses to these four items to form a composite measure of *embarrassment*, consistent with research that treats embarrassment and shame as closely related emotions (e.g., Borg, Staufienbiel, & Scherer, 1988; Carver & Scheier, 1990; Izard, 1977; Kaufman, 1989). The results reported below are all robust to measuring embarrassment using only “embarrassed” and “self-conscious” (correlation coefficient $r = .68, p < .0001$).

We also collected measures to address potential alternative explanations for our findings. First, participants in the middle-school difficulty condition might have felt that the trivia questions were more difficult than they had initially been led to believe and as a result found the challenge unfair, which might account for their decision to disengage from the original set of questions. To examine this possibility, we asked participants to rate the extent to which they found the results they had achieved from the trivia challenge to be unfair (1 = *not at all*; 7 = *very much*).⁹ Second, it is possible that participants in the middle-school difficulty condition felt more negative emotions in general after they worked on questions that were harder than they anticipated, which might have further reduced their

persistence. To tease apart the effect of negative affect from our proposed effect of embarrassment, we asked participants to rate the extent to which they felt angry, upset, frustrated, and disappointed (1 = *not at all*; 7 = *very much*). We averaged these four items to form a composite measure of *negative affect* (Cronbach’s $\alpha = .89$).

Third, another potential concern is that, after performing below their expectations, participants in the middle-school difficulty condition could have had reduced confidence in their ability to answer questions on the original topics relative to participants in the expert difficulty condition. A larger decrease in confidence could potentially cause participants in the middle-school difficulty condition to quit answering questions on the original topics sooner than participants in the expert difficulty condition (Carver & Scheier, 1990). To test this possibility, we asked participants to indicate their confidence at two points in time: before beginning the trivia challenge (participants predicted the percentage of questions they would answer correctly) and after the 20th question (participants predicted the percentage of questions they would answer correctly if they continued to answer questions on the original set of topics). To measure *change in confidence*, we calculated the difference in participants’ confidence ratings between the two measures; negative values indicated that participants became less confident. In addition to *change in confidence*, we also analyzed participants’ absolute level of confidence after the 20th question. Past research has highlighted confidence judgments as a determinant of persistence decisions (Bandura, 1997; Carver et al., 1979; Whyte et al., 1997; Whyte & Saks, 2007). For simplicity, we focus on *change in confidence* in the Results section, but we report analyses related to participants’ absolute levels of confidence in Web Appendix G.

Results

Persistence. Across the first 20 questions, participants in the middle-school difficulty and expert difficulty conditions both performed poorly (answering 22.76% vs. 23.07% of questions correctly, respectively, $p = .80$). Following Hypothesis 1, we predicted that, after exhibiting poor initial performance, people facing high external expectations (i.e., those in the middle-school difficulty condition) would be less persistent than those facing low external expectations (i.e., those in the expert difficulty condition). Indeed, as shown in Figure 4 (Panel A), we found that participants in the middle-school difficulty condition switched

⁹ We measured unfairness using three items (“unfair,” “wrong,” and “legitimate”). However, these three items had a low inter-item reliability (Cronbach’s $\alpha = .61$). We thank the anonymous reviewer who suggested that we focus on the “unfair” item because participants might have misinterpreted the items “legitimate” and “wrong” as referring to whether the trivia challenge provided accurate answers. Our results are robust if we average the three items to create a composite measure of *unfairness*, with “legitimate” reverse-coded.

topics sooner ($M = 22.44$, $SD = 6.34$) than did participants in the expert difficulty condition ($M = 29.89$, $SD = 11.38$), $t(302) = 7.07$, $p < .0001$.¹⁰

Embarrassment. In Hypothesis 2, we predicted that, following poor initial performance, people facing high external expectations would feel more concerned about violating social expectations and more embarrassed than those facing low external expectations. Indeed, as shown in Figure 4 (Panel B), we found that participants in the middle-school difficulty condition felt more embarrassed ($M = 2.65$, $SD = 1.39$) than did participants in the expert difficulty condition ($M = 3.08$, $SD = 1.46$), $t(302) = 2.64$, $p = .009$.

Mediation by embarrassment. Next, we tested whether the effect of external expectations on persistence could be explained by differences in embarrassment (Hypothesis 3). In an OLS regression, embarrassment was a significant and negative predictor of persistence ($\beta = -1.16$, $p < .001$), and the effect of facing high (vs. low) external expectations on persistence was significantly reduced from $\beta = -7.47$ ($p < .001$) to $\beta = -6.95$ ($p < .001$) when embarrassment was included in the model. A 5,000-sample bootstrap analysis estimated an indirect effect of external expectations on persistence via embarrassment as -0.50 (bootstrapped $SE = 0.26$), and the 95% bias-corrected confidence interval of the indirect effect $[-1.17, -0.11]$ did not include 0. Thus, embarrassment mediated the effect of external expectations on persistence, supporting Hypothesis 3.

Alternative explanations. Participants in the middle-school difficulty condition reported feeling (a) that the contest was significantly less fair ($M_{\text{unfairness}} = 2.96$) and (b) that their confidence had decreased to a greater extent ($M_{\text{change in confidence}} = -40.84\%$) than participants in the expert difficulty condition ($M_{\text{unfairness}} = 2.19$), $t(302) = 3.79$, $p = .0001$; ($M_{\text{change in confidence}} = -5.27\%$), $t(302) = 20.01$, $p < .0001$. We found no significant differences

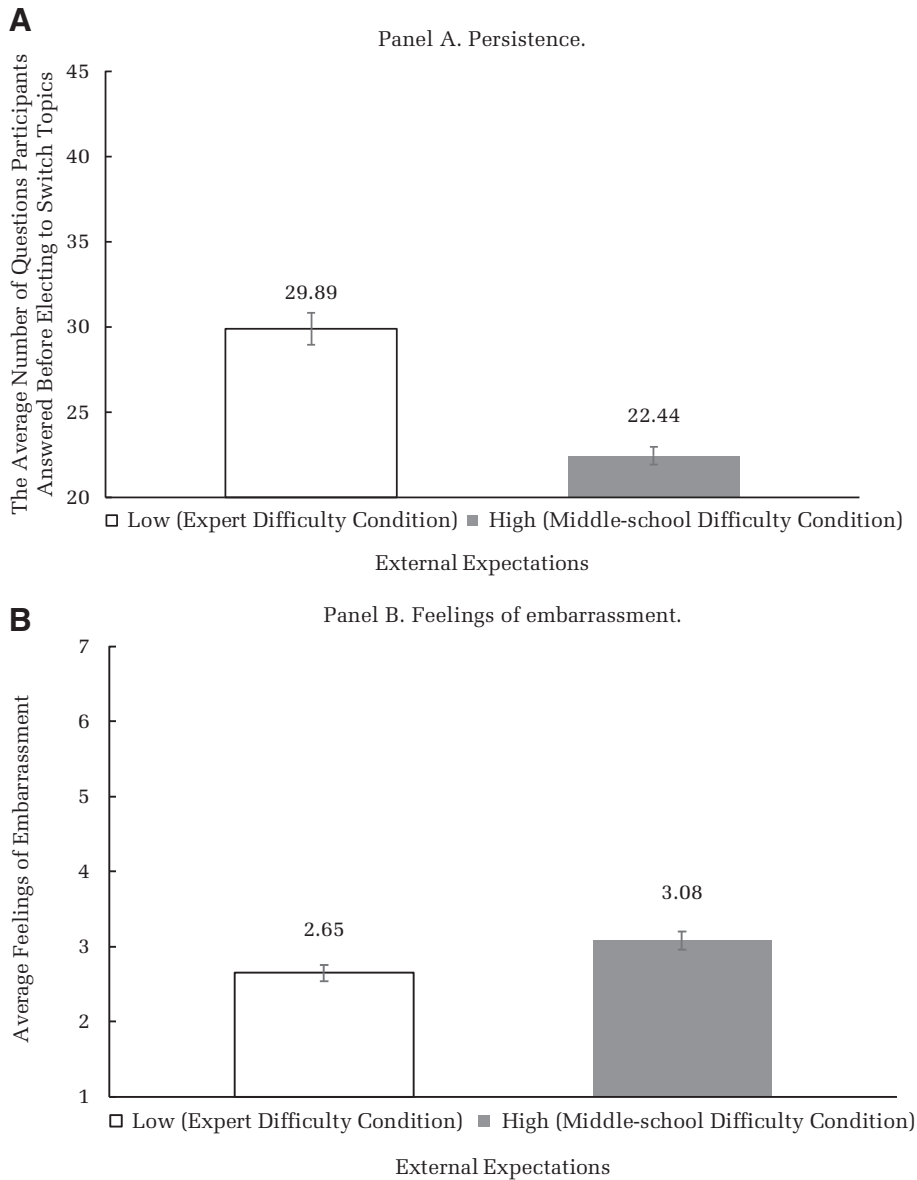
in reported negative affect between participants in the middle-school difficulty ($M_{\text{negative affect}} = 2.86$) and expert difficulty conditions ($M_{\text{negative affect}} = 2.58$) at the 5% level, $t(302) = 1.72$, $p = .09$. We conducted an OLS regression to predict persistence using an indicator for the middle-school difficulty condition, embarrassment, as well as our measures of unfairness, negative affect, and reduced confidence. In this regression, embarrassment remained a significant and negative predictor of persistence ($\beta = -1.96$, $p < .001$). A 5,000-sample bootstrap analysis confirmed that, while controlling for the additional variables measured, the indirect effect of external expectations on persistence via embarrassment was significant (indirect effect = -0.84 ; bootstrapped $SE = 0.39$; the 95% bias-corrected confidence interval was $[-1.79, -0.22]$). Importantly, none of the other variables we measured (unfairness, negative affect, reduced confidence) significantly mediated the effect of external expectations on persistence, regardless of whether these measures were tested independently or simultaneously (see Web Appendix G for details).

Discussion

In this experiment, we measured persistence by assessing how long participants in a trivia challenge continued to answer questions on the same set of topics before electing to switch to a new set of topics for a lower rate of pay. Our design mimicked common situations in organizations in which employees choose to either persist on a project or switch to a different one. This study provides additional support for Hypothesis 1, which posited that, when individuals experience poor initial performance, those facing high external expectations are less persistent than those facing low external expectations. In addition, we found that, following poor initial performance, individuals facing high external expectations experience greater impression management concerns as measured by elevated feelings of embarrassment (Hypothesis 2), which drives them to be less persistent (Hypothesis 3). Furthermore, the results of this study were inconsistent with several alternative explanations for the relationship between poor initial performance, high external expectations, and persistence. The evidence we collected is inconsistent with the possibility that fairness concerns, negative affect, or a reduction in confidence account for our results. Taken together, we found that embarrassment plays a unique role

¹⁰ In light of the right-censoring nature of our data, we conducted a robustness check using a survival analysis. Specifically, we ran a proportional hazards regression, specified that the data were right censored at 45, and defined the event of interest to be switching topics prior to the 45th trivia question. A positive coefficient estimated in this model on the indicator for our middle-school difficulty condition ($\beta = 2.05$, $p < .001$) indicated that the hazard rate of quitting in the middle-school difficulty condition is 7.77 times (i.e., $\exp(2.05)$) as large as that in the expert difficulty condition.

FIGURE 4
Persistence and Feelings of Embarrassment as a Function of External Expectations in Study 2



Note: Error bars represent the standard error of the mean.

in prompting individuals facing high external expectations to quit sooner than others when facing adversity.

In general, high expectations and high self-efficacy increase motivation and improve performance (Bandura, 1997; Eden, 1990; Vroom, 1964). In our experiment, we manipulated performance expectations, but did not find a link between expectations and performance. We suspect that the nature of our task, a trivia challenge, limited the effects of motivation. That is, greater effort is likely to improve performance

across a variety of domains, but it is unlikely to improve the ability to recall trivia.

GENERAL DISCUSSION

Across a field study and a laboratory study, we investigated how external performance expectations influence persistence decisions. In Study 1, we analyzed more than 30 years of archival data from professional men’s tennis matches. Using a quasi-experimental RD approach, we found that,

following poor initial performance (a loss in the first set), competitors who were expected to win were less persistent (i.e., they were more likely to quit mid-way through a match) than similarly talented competitors who were expected to lose. In Study 2, we manipulated external performance expectations and found the same pattern of results: participants who faced higher performance expectations were less persistent than those who faced lower performance expectations following poor initial performance. In addition, we found that high external expectations increased impression management concerns and triggered embarrassment, which mediated the combined effect of poor performance and high external expectations on persistence. The combination of field and laboratory data provides convergent evidence for both the external and internal validity of the negative link between expectations and persistence in the face of adversity.

Theoretical and Practical Implications

Our research makes several important contributions to management theory. First, within organizations, managers and employees constantly face decisions about whether to persist in the face of adversity or redirect their effort. In spite of the enormous importance these decisions can have, there are significant gaps in our understanding of when and why individuals persist. We demonstrate that the decision to persist can be significantly influenced by psychological factors that have little to do with the economic costs and benefits of persisting. Although individuals' confidence in their ability to overcome initial setbacks heavily influences persistence decisions, as highlighted by past research (e.g., Bandura, 1997; Berger & Pope, 2011; Carver et al., 1979; Whyte et al., 1997; Whyte & Saks, 2007), we identify impression management considerations, or, more precisely, the desire to avoid embarrassment, as another important antecedent of quitting that, in some settings, may trump the influence of confidence and cause individuals with high external expectations to be less persistent.

Second, in contrast to the extant literature that has overwhelmingly highlighted the advantages of being expected to perform well (see Eden, 1990, and McNatt, 2000, for meta-analyses), we both theorize and demonstrate that facing high external expectations can, in the face of early setbacks, be a burden. Following poor initial performance, high external expectations amplify impression management concerns producing feelings of embarrassment and causing people to quit more often than they

would if they had faced low external expectations, even when quitting entails forgoing opportunities to earn substantial rewards (e.g., substantial prize money in the case of Study 1). Our theorizing and findings in no way contradict the robust finding from past research that high external expectations, on average, lead to better performance. Instead, we note that performance depends on factors beyond motivation such as luck and extenuating circumstances. As a result, people with high expectations may not always perform well. We fill a gap in the organizational literature by demonstrating that high expectations can lower persistence when initial performance is poor.

Third, our findings add to a small but growing collection of field studies that link psychological factors to unethical behaviors (Gino & Pierce, 2010; Greenberg, 1993; Kilduff, Galinsky, Gallo, & Reade, 2016). According to ATP rules, players are only allowed to quit mid-match due to an injury. Given the sensitivity of quitting decision to the combination of high expectations and poor initial performance, our findings suggest that some claims of physical injury are likely illegitimate. We call for future research to examine the relationship between performance expectations and ethical decisions in field settings.

Our findings also offer a number of prescriptions for managers. First, we shed light on the importance of managing subordinates' performance expectations. Past research suggests that managers can boost employees' motivation and job performance by communicating positive expectations (McNatt, 2000). However, our findings reveal that conveying positive expectations may come at a cost under some circumstances. If subordinates face high expectations, but experience an early setback, they may feel particularly concerned about their public image, experience embarrassment, and reduce persistence. In practice, this pattern of behavior may take many forms, such as stepping down from a leadership position, quitting a competition, decreasing involvement in a project, or reallocating effort from one task to another. Regardless of its specific form, failures to persist in the workplace can be costly to both individuals and organizations. Managers' awareness of the relationship between expectations and persistence may be particularly important in competitive fields that attract high performers. For example, candidates who enter top-tier consulting or banking jobs, students who enter elite colleges, and faculty who begin careers at top-tier institutions may be particularly prone to quitting when they encounter setbacks. To reduce the likelihood of premature

quitting and the harmful effects of embarrassment, managers should focus attention on employees who face high expectations and experience early setbacks and develop strategies to help them. For example, it may be helpful to foster a culture that accepts and learns from failure and to cultivate a growth mindset (Dweck, 2006).

Our findings from Study 1 also highlight a harmful consequence of ranking systems. Greater access to data and analytical tools has facilitated increased reliance on ranking systems in many industries (Mills & Mills, 2014). Ranking systems communicate external expectations and may therefore create impression management anxiety for employees facing high expectations. Our findings suggest that ranking systems should be used with caution and add to an ongoing discussion about the benefits and costs of ranking employees (Barankay, 2012; Charness, Masclet, & Villeval, 2014) and making rankings transparent (Bernstein & Li, 2016; Song, Tucker, Murrell, & Vinson, 2018).

Limitations and Future Directions

We combined analyses of laboratory and archival data to establish the robustness of our findings and their relevance to management practice. The sports context we studied represents a large, complex organization with many stakeholders and interest groups (e.g., players, coaches, sponsors, spectators, and the media) who are deeply invested in the organization. Also, there are many similarities between the organization of men's professional tennis and more traditional corporations. For example, tennis players and businesspeople are rewarded based on their performance, subject to impression management concerns, and accountable to stakeholders (e.g., sponsors vs. investors). Management scholars regularly use sports data to study organizational phenomena, leveraging the availability of detailed and objective performance metrics (e.g., Kilduff, Elfenbein, & Staw, 2010; Larrick, Timmerman, Carton, & Abrevaya, 2011; Marr & Thau, 2014) as we have done here. Still, we call for future research to explore the relationship between expectations and persistence in more traditional organizational environments.

One question that emerges from this investigation is whether the expectation–persistence relationship documented here represents a burden associated with high external expectations or a benefit of low external expectations. One interpretation of our results is that individuals facing low expectations have little to lose

from a failure and much to gain from a success and thus are more persistent in the face of adversity, even when their odds of success are low. In this work, we have focused on documenting the novel, negative expectation–persistence relationship. Future research that differentiates the psychology of facing high versus low expectations could advance our growing understanding of the interpersonal and motivational benefits of being an underdog (e.g., Lount et al., 2017; Nurmohamed, 2014; Paharia, Keinan, Avery, & Schor, 2011).

Both of our studies are characterized by high performance visibility. In Study 1, individuals performed in public, and, in Study 2, individuals knew that their performance would become public. This feature of our studies highlights the influence of external performance expectations and impression management concerns, and mirrors the constant, high visibility of job performance in roles ranging from sales, to trading, to call center operation. Across many organizational contexts, performance is highly visible (e.g., because performance rankings are posted, periodic updates are provided on product development, contests are conducted, etc.). In other contexts (e.g., promotions, elections), outcomes are more visible than interim performance. Future work should explore the relationship between expectations and persistence when interim performance and even final performance outcomes are invisible to others. In these settings, public image concerns should exert far less influence on behavior (Leary & Kowalski, 1990), and high external expectations may not reduce persistence. When performance is difficult to observe, internal expectations (the expectations that people have for themselves) are likely to matter far more than external expectations. Investigating whether or not high internal expectations lead people to reduce their persistence following poor initial performance is an interesting future research direction. On the one hand, individuals with high internal expectations may perceive poor initial performance as personally disappointing and may thus adopt an exit strategy in order to maintain a positive self-image. On the other hand, self-image maintenance often “occurs entirely on a cognitive level” (Leary & Kowalski, 1990: 35) and may not elicit behavioral responses such as quitting.

Prior research on quitting has focused primarily on its downside (such as reduced or eliminated chances of success and the social stigma associated with quitting; e.g., Eriksson, Mao, & Villeval, 2017; Fershtman & Gneezy, 2011) and overlooked impression management benefits that can, and often do, outweigh the costs of quitting. By quitting, people can withdraw from

situations that produce embarrassment, avoid being held accountable for poor outcomes, and take the opportunity to rebuild their public image by excelling elsewhere. In particular, when attributions for poor performance are ambiguous and when the prospect of an initiative is uncertain, quitting with a plausibly valid excuse can not only provide a reasonable explanation for poor performance, but also preserve the impression that the “quitter” might have succeeded if not for extenuating circumstances (e.g., family obligations, health issues). Potentially interesting avenues for future research include exploring how people balance the impression management benefits and costs of quitting and how the perceived legitimacy of a quitting excuse moderates the relationship we document between external expectations and persistence.

Extending prior work that has identified high performance expectations as an asset, we reveal that, in the face of adversity, high expectations can also be a liability when it comes to persistence. Our field and laboratory studies together demonstrate that, when the going gets tough, high external expectations can lead people to experience greater embarrassment and reduce their persistence. We encourage future research to further our understanding of both the benefits and the burdens of high expectations.

REFERENCES

- Angrist, J. D., & Pischke, J. S. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- ATP Official Rulebook. (2018). Retrieved from <https://www.atpworldtour.com/en/corporate/rulebook>.
- Babad, E. Y., Inbar, J., & Rosenthal, R. 1982. Pygmalion, Galatea, and the Golem: Investigations of biased and unbiased teachers. *Journal of Education & Psychology*, 74: 459–474.
- Bandura, A. 1982. Self-efficacy mechanism in human agency. *The American Psychologist*, 37: 122–147.
- Bandura, A. 1997. *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Barankay, I. 2012. *Rank incentives: Evidence from a randomized workplace experiment*. Working paper. Philadelphia, PA: Wharton Business School.
- Baumeister, R. F., Hamilton, J. C., & Tice, D. M. 1985. Public versus private expectancy of success: Confidence booster or performance pressure? *Journal of Personality and Social Psychology*, 48: 1447–1457.
- Baumeister, R. F., & Steinhilber, A. 1984. Paradoxical effects of supportive audiences on performance under pressure: The home field disadvantage in sports championships. *Journal of Personality and Social Psychology*, 47: 85–93.
- Berger, J., & Pope, D. 2011. Can losing lead to winning? *Management Science*, 57: 817–827.
- Berglas, S., & Jones, E. E. 1978. Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement. *Personality and Social Psychology Bulletin*, 4: 200–206.
- Bernstein, E. S., & Li, S. X. 2016. *Seeing where you stand: From performance feedback to performance transparency*. Working paper. Boston, MA: Harvard Business School.
- Borg, I., Staufenbiel, T., & Scherer, K. R. 1988. On the symbolic basis of shame. In K. R. Scherer (Ed.), *Facets of emotion: Recent research*: 79–98. Hillsdale, NJ: Erlbaum.
- Breznik, K., & Batagelj, V. 2012. Retired matches among male professional tennis players. *Journal of Sports Science & Medicine*, 11: 270–278.
- Brooks, A. W., & Schweitzer, M. E. 2011. Can nervous Nelly negotiate? How anxiety causes negotiators to make low first offers, exit early, and earn less profit. *Organizational Behavior and Human Decision Processes*, 115: 43–54.
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. 2008. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90: 414–427.
- Carver, C. S., Blaney, P. H., & Scheier, M. F. 1979. Reassertion and giving up: The interactive role of self-directed attention and outcome expectancy. *Journal of Personality and Social Psychology*, 37: 1859–1870.
- Carver, C. S., & Scheier, M. F. 1990. Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97: 19–35.
- Charness, G., Masclet, D., & Villeval, M. C. 2014. The dark side of competition for status. *Management Science*, 60: 38–55.
- Chen, H., Ham, S. H., & Lim, N. 2011. Designing multi-person tournaments with asymmetric contestants: An experimental study. *Management Science*, 57: 864–883.
- Cheryan, S., & Bodenhausen, G. V. 2000. When positive stereotypes threaten intellectual performance: The psychological hazards of “model minority” status. *Psychological Science*, 11: 399–402.
- Conlon, E. J. 1980. Feedback about personal and organizational outcomes and its effect on persistence of planned behavioral changes. *Academy of Management Journal*, 23: 267–286.

- Davidson, O. B., & Eden, D. 2000. Remedial self-fulfilling prophecy: Two field experiments to prevent Golem effects among disadvantaged women. *Journal of Applied Psychology*, 85: 386–398.
- Dong, P., Huang, X., & Wyer, R. S., Jr. 2013. The illusion of saving face: How people symbolically cope with embarrassment. *Psychological Science*, 24: 2005–2012.
- Dweck, C. S. 2006. *Mindset: The new psychology of success*. New York, NY: Random House.
- Eden, D. 1984. Self-fulfilling prophecy as a management tool: Harnessing Pygmalion. *Academy of Management Review*, 9: 64–73.
- Eden, D. 1990. *Pygmalion in management: Productivity as self-fulfilling prophecy*. Lexington, MA: Lexington Books.
- Eden, D., & Aviram, A. 1993. Self-efficacy training to speed reemployment: Helping people to help themselves. *Journal of Applied Psychology*, 78: 352–360.
- Eriksson, T., Mao, L., & Villeval, M. C. 2017. Saving face and group identity. *Experimental Economics*, 20: 622–647.
- Fershtman, C., & Gneezy, U. 2011. The tradeoff between performance and quitting in high-power tournaments. *Journal of the European Economic Association*, 9: 318–336.
- Flammer, C. 2015. Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science*, 61: 2549–2568.
- Gelman, A., & Imbens, G. 2014. *Why high-order polynomials should not be used in regression discontinuity designs*. Working paper no. 20405. Cambridge, MA: National Bureau of Economic Research.
- Gibson, B., Sachau, D. A., Doll, B., & Shumate, R. 2002. Sandbagging in competition: Responding to the pressure of being the favorite. *Personality and Social Psychology Bulletin*, 26: 56–70.
- Gimeno, J., Folta, R. B., Cooper, A. C., & Woo, C. Y. 1997. Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly*, 42: 750–783.
- Gino, F., & Pierce, L. 2010. Robin Hood under the hood: Wealth-based discrimination in illicit customer help. *Organization Science*, 21: 1176–1194.
- Girard, O., Eicher, F., Fourchet, F., Micallef, J. P., & Millet, G. P. 2007. Effects of the playing surface on plantar pressures and potential injuries in tennis. *British Journal of Sports Medicine*, 41: 733–738.
- Green, S. G., Welsh, M. A., & Dehler, G. E. 2003. Advocacy, performance, and threshold influences on decisions to terminate new product development. *Academy of Management Journal*, 46: 419–434.
- Greenberg, J. 1993. Stealing in the name of justice: Informational and interpersonal moderators of theft reactions to underpayment inequity. *Organizational Behavior and Human Decision Processes*, 54: 81–103.
- Greenberg, J., Pyszczynski, T., & Paisley, C. 1984. Effect of extrinsic incentives on use of test anxiety as an anticipatory attributional defense: Playing it cool when the stakes are high. *Journal of Personality and Social Psychology*, 47: 1136–1145.
- Hahn, J., Todd, P., & van der Klaauw, W. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69: 201–209.
- Imbens, G., & Lemieux, T. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142: 615–635.
- Imbens, G., & Kalyanaraman, K. 2009. *Optimal bandwidth choice for the regression discontinuity estimator*. Working paper no. 14726. Cambridge, MA: National Bureau of Economic Research.
- Imbens, G., & Kalyanaraman, K. 2012. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79: 933–959.
- Isaac, M. S., & Schindler, R. M. 2014. The top-ten effect: Consumer's subjective categorization of ranked lists. *The Journal of Consumer Research*, 40: 1181–1202.
- Izard, C. E. 1977. *Human emotions*. New York, NY: Plenum.
- Jensen, M., & Kim, H. 2015. The real Oscar curse: The negative consequences of positive status shifts. *Organization Science*, 26: 1–21.
- Jensen, M., Kim, H., & Kim, B. 2012. Meeting expectations: A role-theoretic perspective on reputation. In M. L. Barnett & T. G. Pollock (Eds.), *The Oxford handbook of corporate reputation*: 140–159. Oxford, U.K.: Oxford University Press.
- Kaufman, G. 1989. *The psychology of shame: Theory and treatment of shame-based syndromes*. New York, NY: Springer.
- Keltner, D., & Buswell, B. N. 1997. Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin*, 122: 250–270.
- Kilduff, G. J., Elfenbein, H. A., & Staw, B. M. 2010. The psychology of rivalry: A relationally dependent analysis of competition. *Academy of Management Journal*, 53: 943–969.
- Kilduff, G. J., Galinsky, A. D., Gallo, E., & Reade, J. J. 2016. Whatever it takes to win: Rivalry increases unethical behavior. *Academy of Management Journal*, 59: 1508–1534.

- Kolditz, T. A., & Arkin, R. M. 1982. An impression management interpretation of the self-handicapping strategy. *Journal of Personality and Social Psychology*, 43: 492–502.
- Kovács, B., & Sharkey, A. J. 2014. The paradox of publicity: How awards can negatively affect the evaluation of quality. *Administrative Science Quarterly*, 59: 1–33.
- Larrick, R. P., Timmerman, T. A., Carton, A. M., & Abrevaya, J. 2011. Temper, temperature, and temptation: Heat-related retaliation in baseball. *Psychological Science*, 22: 423–428.
- Leary, M. R., & Kowalski, R. M. 1990. Impression management: A literature review and two-component model. *Psychological Bulletin*, 107: 34–47.
- Lee, D. S., & Lemieux, T. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48: 281–355.
- Lim, N. 2010. Social loss aversion and optimal contest design. *Journal of Marketing Research*, 47: 777–787.
- Lount, R. B., Jr., Pettit, N. C., & Doyle, S. P. 2017. Motivating underdogs and favorites. *Organizational Behavior and Human Decision Processes*, 141: 82–93.
- Luca, M. 2014. *Reviews, reputation, and revenue: The case of Yelp.com*. Working paper no. 12-016. Boston, MA: Harvard Business School.
- Luca, M., & Smith, J. 2013. Salience in quality disclosure: Evidence from the U.S. News college rankings. *Journal of Economics & Management Strategy*, 22: 58–77.
- March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2: 71–87.
- Marr, J. C., & Thau, S. 2014. Falling from great (and not-so-great) heights: How initial status position influences performance after status loss. *Academy of Management Journal*, 57: 223–248.
- McNatt, D. B. 2000. Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, 85: 314–322.
- Miller, R. S. 1992. The nature and severity of self-reported embarrassing circumstances. *Personality and Social Psychology Bulletin*, 18: 190–198.
- Miller, R. S., & Leary, M. R. 1992. Social sources and interactive functions of embarrassment. In M. Clark (Ed.), *Emotion and social behavior*: 322–329. New York, NY: Russell Sage Foundation.
- Mills, M. P., & Mills, M. A. 2014, 29 June. Every datum tells a story: The dawning of the age of meta-information. *City Journal*. Retrieved from <http://www.city-journal.org/html/every-datum-tells-story-11392.html>
- Mishina, Y., Dykes, B. J., Block, E. S., & Pollock, T. G. 2010. Why “good” firms do bad things: The effects of high aspirations, high expectations, and prominence on the incidence of corporate illegality. *Academy of Management Journal*, 53: 701–722.
- Modigliani, A. 1971. Embarrassment, facework, and eye contact: Testing a theory of embarrassment. *Journal of Personality and Social Psychology*, 17: 15–24.
- Mosher, D. L., & White, B. B. 1981. On differentiating shame and shyness. *Motivation and Emotion*, 5: 61–74.
- Nurmohamed, S. 2014. Over or under? *The motivational implications of an underdog image*. Unpublished doctoral dissertation. University of Michigan, Ann Arbor, Michigan.
- Paharia, N., Keinan, A., Avery, J., & Schor, J. B. 2011. The underdog effect: The marketing of disadvantage and determination through brand biography. *The Journal of Consumer Research*, 37: 775–790.
- Petkova, A. P., Wadhwa, A., Yao, X., & Jain, S. 2014. Reputation and decision making under ambiguity: A study of U.S. venture capital firms’ investments in the emerging clean energy sector. *Academy of Management Journal*, 57: 422–448.
- Pettit, N. C., Sivanathan, N., Gladstone, E., & Marr, J. C. 2013. Rising stars and sinking ships: Consequences of status momentum. *Psychological Science*, 24: 1579–1584.
- Pierce, L., Dahl, M. S., & Nielsen, J. 2013. In sickness and in wealth: Psychological and sexual costs of income comparison in marriage. *Personality and Social Psychology Bulletin*, 39: 359–374.
- Pope, D. 2009. Reacting to rankings: Evidence from “America’s Best Hospitals.” *Journal of Health Economics*, 28: 1154–1165.
- Rosenthal, R., & Jacobson, L. 1968. *Pygmalion in the classroom: Teacher expectation and pupils’ intellectual development*. New York, NY: Holt, Rinehart & Winston.
- Song, H., Tucker, A. L., Murrell, K. L., & Vinson, D. R. 2018. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science*, 64: 2628–2649.
- Steele, C. M., & Aronson, J. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69: 797–811.
- Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. 1996. Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, 70: 1256–1269.
- Vroom, V. H. 1964. *Work and motivation*. New York, NY: Wiley.

- Weber, M., & Camerer, C. F. 1998. The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization*, 33: 167–184.
- Whyte, G., & Saks, A. M. 2007. The effects of self-efficacy on behavior in escalation situations. *Human Performance*, 20: 23–42.
- Whyte, G., Saks, A. M., & Hook, S. 1997. When success breeds failure: The role of self-efficacy in escalating commitment to a losing course of action. *Journal of Organizational Behavior*, 18: 415–432.
- Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. Boston, MA: The MIT Press.
- Zhang, D. J., Allon, G., & Van Mieghem, J. A. 2017. Does social interaction improve learning outcomes? Evidence from field experiments on massive open online courses. *Manufacturing & Service Operations Management*, 19: 347–367.



Hengchen Dai (hengchen.dai@anderson.ucla.edu) is an assistant professor of management and behavioral decision making at Anderson School of Management at University of California, Los Angeles. She received her PhD from the Wharton School, University of Pennsylvania. Her research

focuses on motivation, decision making, employee well-being, and social influence.

Berkeley Dietvorst (berkeley.dietvorst@chicagobooth.edu) is an assistant professor of marketing at the University of Chicago's Booth School of Business. He received his PhD from the Wharton School, University of Pennsylvania. His research focuses on judgment and decision making, with a particular interest in consumers' and managers' use of decision aids.

Bradford Tuckfield (bradford.tuckfield@aexp.com) is a senior data scientist at American Express GBT. He received his PhD from the Wharton School, University of Pennsylvania.

Katherine L Milkman (kmilkman@wharton.upenn.edu) is a professor of operations, information and decisions and the Evan C. Thompson Endowed Term Chair for Excellence in Teaching at the Wharton School at the University of Pennsylvania. Her research focuses on behavioral economics, understanding self-control failures, and improving decision making.

Maurice Schweitzer (schweitzer@wharton.upenn.edu) is the Cecilia Yen Koo professor of operations, information and decisions at the Wharton School at the University of Pennsylvania. His research focuses on emotions, decision making, and the negotiation process.



Copyright of Academy of Management Journal is the property of Academy of Management and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.