QUITTING: THE DOWNSIDE OF HIGH PERFORMANCE EXPECTATIONS

Hengchen Dai[1], Berkeley J. Dietvorst[2], Bradford Tuckfield[3]
Katherine L. Milkman[4], Maurice E. Schweitzer[4]

[1] Olin Business School, Washington University in St. Louis
[2] Booth School of Business, University of Chicago
[3] Berkadia
[4] The Wharton School, University of Pennsylvania

QUITTING: THE DOWNSIDE OF HIGH PERFORMANCE EXPECTATIONS

Abstract

The benefits of high performance expectations are well documented. We identify a novel *disadvantage* of high performance expectations: Compared to individuals with low performance expectations, individuals with high performance expectations are less likely to persist in the face of adversity and have greater image management concerns. In a field study of 328,515 men's professional tennis matches (Study 1), we employ a regression discontinuity design to demonstrate that favorites are significantly more likely to quit than are underdogs after losing the first set of the match. We replicate this pattern of results in a laboratory experiment (Study 2); individuals with high performance expectations are less persistent following early disappointments than those with low performance expectations. In another laboratory experiment (Study 3), we show that following poor initial performance, individuals with high expectations experience greater embarrassment and shame than those with low expectations. Taken together, we find that when the going gets tough, having high performance expectations can cause people to be less persistent. We do not observe this pattern, however, when initial performance is strong (Studies 1 and 3). We discuss organizational and managerial implications of our findings.

QUITTING: THE DOWNSIDE OF HIGH PERFORMANCE EXPECTATIONS

Executives, managers, and employees constantly face decisions to either persist in their current endeavor or redirect their effort. For example, after learning new information, entrepreneurs decide to persist or abandon their venture (Gimeno, Folta, Cooper, & Woo, 1997), project managers decide to persist or terminate the project (Green, Welsh, & Dehler, 2003), and employees decide whether or not to persist on a current task or redirect their effort (Conlon, 1980). The decision to persist or quit can have profound implications. Persisting in a failing venture can be ruinous, but quitting prematurely means forgoing opportunities to reap substantial rewards (e.g., Brooks & Schweitzer, 2011; March, 1991; Weber & Camerer, 1998). In this work, we investigate the decision to persist and focus on an important antecedent: performance expectations.

Extant research has described significant advantages for individuals who hold high performance expectations. In fact, several influential theoretical frameworks, including self-fulfilling prophecy theory (Eden, 1990, 1992; McNatt, 2000; McNatt & Judge, 2004; Merton, 1948; Rosenthal & Jacobson, 1968a, 1968b) and self-efficacy theory (Bandura, 1982, 1997) postulate that individuals who believe that they will succeed are more likely to succeed. This substantial stream of research has concluded that high expectations help individuals succeed, because they take actions that are consistent with high positive expectations and because they are more resilient in the face of adversity (Bandura, 1982, 1997; Berger & Pope, 2011; Eden & Aviram, 1993).

We challenge prior work by demonstrating that high performance expectations can *harm* persistence. In our research, we identify a critical factor that accompanies high performance expectations and has been overlooked in the study of performance expectations and persistence:

image management motives, the desire to maintain both a positive self-image and a positive public image (Finch & Cialdini, 1989). We reveal that individuals with high performance expectations hold stronger self-image and public image concerns than individuals with low performance expectations. This difference in image concerns matters when individuals experience poor performance. Individuals with high performance expectations perceive poor performance as a violation of personal and external expectations and thus feel more embarrassed and ashamed. In contrast, individuals with low expectations perceive objectively identical (poor) performance as consistent with their expectations and feel less embarrassed and ashamed.

Even though individuals with high performance expectations may be more confident in their ability to succeed than individuals with low performance expectations, feelings of embarrassment and shame may cause them to be less persistent following poor performance. For those with high performance expectations, shifting resources away from the focal initiative will decrease exposure to the situation that gives rise to embarrassment and shame, and diminish their accountability for further poor performance. Further, by quitting or reducing persistence and citing an excuse for their poor performance, individuals can provide an alternative account, separate from incompetence, for their disappointing initial showing.

We test our theory across three studies, including a field study analyzing over 30-years of data on professional men’s tennis matches and two laboratory experiments. We find that, compared to individuals with low performance expectations, those with high expectations feel more embarrassed and ashamed after exhibiting poor initial performance and are less likely to persist. Further, we find that when initial performance is strong, high performance expectations do not elicit feelings of embarrassment or shame or reduce persistence.

Our research makes several important theoretical contributions. First, we provide insight into how people make persistence decisions by identifying an often overlooked motive that undermines persistence. 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 the conclusions of existing research that assumes people make persistence decisions based solely on their chances of success (e.g., Bandura, 1997).

Second, our research challenges prior research that has focused on the benefits of holding high performance expectations (e.g., Davidson & Eden, 2000; McNatt, 2000; McNatt & Judge, 2004). By documenting a novel drawback to holding high expectations, our research addresses a recent call for research to explore the downsides of positive expectations (George, Dahlander, Graffin, & Sim, 2016). Taken together, our findings advance our understanding of persistence and the psychological processes that are shaped by performance expectations.

Performance Expectations and Persistence Decisions

Individuals within organizations face both self-imposed (Bandura, 1986: 391; McNatt & Judge, 2004) and externally-imposed performance expectations (e.g., from supervisors, coworkers, clients, and the media; Eden, 1990; McNatt, 2000). Both self-imposed and externally-imposed performance expectations develop from prior experience (Bandura, 1997; Mishina, Dykes, Block, & Pollock, 2010), rankings (Legge & Schmid, 2014; Luca & Smith, 2013; Pope, 2009), reputation (Jensen, Kim, & Kim, 2012; Petkova, Wadhwa, Yao, & Jain, 2014), or affiliation with a high (or low) status group (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003; Kray, Thompson, & Galinsky, 2001; Steele, 1997; Steele & Aronson, 1995). We conceptualize performance expectations as individuals' forecasts of how effectively they will perform on a given task or in a given role, and we conceptualize performance expectations as the combination of self-imposed and externally-imposed expectations.

Past research has identified high expectations as an antecedent of high performance, and low expectations as an antecedent of low performance. Self-fulfilling prophecy theory (Eden, 1990, 1992; McNatt, 2000; McNatt & Judge, 2004; Merton, 1948; Rosenthal & Jacobson, 1968a, 1968b) predicts that people will behave and perform in ways that are consistent with their expectations. Similarly, self-efficacy theory (Bandura, 1982) proposes that people who are more confident in their ability to achieve a performance goal will expend greater effort and work more persistently, which will ultimately lead to superior performance. A related literature on stereotype threat suggests that low expectations caused by the activation of negative, self-relevant stereotypes in a performance domain may reduce effort and lower performance (Cadinu et al., 2003; Kray et al., 2001; Steele, 1997; Steele & Aronson, 1995). A number of empirical studies offer support for these theories across diverse domains (see McNatt, 2000 for a meta-analysis), including in corporations (Chen & Klimoski, 2003; McNatt & Judge, 2004), the military (Eden & Zuk, 1995; Schrank, 1968), academia (Babad, Inbar, & Rothal, 1982; Chapman and McCauley, 1993; Oz & Eden, 1994), and experimental laboratories (Gold, 1999; Steele & Aronson, 1995).

One particular mechanism through which high expectations produce better outcomes is by helping people to cope with poor initial performance and persist in the face of adversity. Self-efficacy theory states that individuals with high self-efficacy are more likely to sustain or even heighten effort in the face of setbacks than those with low self-efficacy (Bandura, 1982, 1997). Consistent with this theory, people with high self-efficacy have been shown to be 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 insulated from the negative effects of initial impasse on their future willingness to collaborate (O'Connor & Arnold, 2001). Even when persistence is not economically rational, individuals high in self-efficacy may believe they can reverse their failing fortunes, and are hence more likely to persist even in a failing enterprise (Whyte, Saks, & Hook, 1997; Whyte & Saks, 2007). Taken together, past research suggests that when faced with poor initial performance, individuals with high expectations are more likely to sustain or even increase their effort because they are more confident in their ability to overcome obstacles than individuals with low performance expectations. In our investigation, we challenge prior research and demonstrate that high performance expectations can harm persistence. We identify an important, and previously neglected, factor that causes people with high performance expectations to be less likely to persist in the face of adversity.

The Role of Image Management

We postulate that image management concerns can cause people with high performance expectations to be less persistent when they face adversity. We use the term image management to refer to both *self-image* management—the attempt to live up to personal standards and maintain a favorable self-image (e.g., Steele, 1988; Swann & Hill, 1982)—and *impression* management—the attempt to meet social expectations and make a good impression on others (Apsler 1975; Leary & Kowalski, 1990; Modigliani 1971; Stevens & Kristof, 1995).

Expectations, image management concerns, and negative, self-conscious emotions. Expectations can have a strong effect on people’s concerns about their public and self-images. Compared to individuals with low performance expectations, individuals with high performance expectations experience greater performance pressure, and subject their performance outcomes to stricter standards (Gibson, Sachau, Doll, & Shumate, 2002; Pettit, Sivanathan, Gladstone, & Marr, 2013). For individuals with high performance expectations, poor performance violates their personal and external standards, but for individuals with low performance expectations, poor performance merely conforms with their low personal and external expectations. Conversely, for individuals with high performance expectations, high performance merely meets their personal and external standards, but for individuals with low performance expectations, high performance exceeds their low personal and external expectations. This logic helps to explain Chen, Ham, and Lim's (2011) finding that individuals with high expectations to win a contest anticipate larger psychological losses from losing the contest and smaller psychological gains from winning than do individuals with low expectations.

Consequently, we expect people with high performance expectations to feel stronger negative emotions following poor performance than those with low performance expectations. Consistent with this prediction, Weinberg, Yukelson, and Jackson (1980) found that how people reacted to a defeat in a physical contest was affected by whether or not they expected to win: participants were more upset when they were told that they had lost to a recovering surgery patient than when they were told that they had been defeated by an athlete. Similarly, Marr and Thau (2014) found that after a status loss, individuals with high initial status experience more self-threat (threat to a central view of the self) than individuals with low initial status. Even outside observers expect that individuals with high performance expectations in a contest will feel more concerned about their public image than individuals with low performance expectations should they lose the contest (Gibson et al., 2002).

In this work, we examine individuals' reactions to poor initial performance in the field and in the laboratory. We examine persistence decisions and focus on how these may be impacted by embarrassment and shame—self-conscious emotions that are associated with violations of expectations (Tangney, 1999). Scholars have proposed that embarrassment and shame are similar but distinct emotions; they result from violations of different types of standards and respond to private and public environments differently (Tangney, Miller, Flicker, & Barlow, 1996). Embarrassment occurs in the company of others (Modigliani, 1971; Tangney et al., 1996) when undesirable events threaten the public image that individuals hope to maintain (Miller, 1992) and when individuals are concerned about what others are thinking about them (Miller & Leary, 1992). In contrast, shame involves changes in self-evaluations and may occur in both public and private performance contexts (Miller & Tangney, 1994).

Both embarrassment and shame are negative, self-conscious emotions that arise when behaviors are incongruent with personal or social standards and when self- or public images are shaken (Tangney, 1999). We propose that when initial performance is poor, individuals with high performance expectations will be more concerned about their self- and public image, experiencing stronger embarrassment and shame, than individuals with low performance expectations. However, we do not expect this difference to emerge when initial performance is strong: strong initial performance is unlikely to elicit negative feelings regardless of individuals' prior expectations.

Next, we propose that feelings of embarrassment and shame experienced by individuals with high performance expectations may lead them to reduce their persistence following poor initial performance as an image management strategy.

Expectations and reducing persistence as an image management strategy. We propose that differences in image concerns (caused by different performance expectations) may influence people's likelihood of decreasing their persistence. Past research suggests that people strive to meet the positive expectations associated with their social or professional roles (Leary & Kowalski, 1990), often by engaging in image management strategies (Jones, Gergen, & Jones, 1963; Leary, Barnes, & Griebel, 1986). The ideal way for people to maintain a positive self- and public image is to actually perform well. However, when good performance seems unattainable, one alternative approach to manage impressions is to change course by reducing persistence. In organizations, reducing persistence can be as explicit as stepping down from an executive position, quitting a high-stakes competition, exiting a negotiation, or terminating a project. It can also take subtler forms, such as switching tasks and reallocating attention and effort between projects. We predict that when an initiative (e.g., a task, project, contest, negotiation) is not proceeding well, individuals with high performance expectations may be more likely to reduce persistence on that initiative than individuals with low performance expectations, for several reasons.

First, if poor initial performance leads people with high expectations, compared to people with low expectations, to feel more embarrassed and ashamed, as we hypothesize, then they may be more likely to quit in order to alleviate these negative self-conscious feelings (Apsler, 1975; Frijda, Kuipers, & ter Schure, 1989). Switching to a different task allows individuals with high expectations to avoid exposure to the source of embarrassment and shame, and may even afford them a clean slate to validate themselves in another area. This prediction is consistent with past research showing that embarrassment and shame can motivate individuals to (a) avoid the conditions that elicit these negative feelings (de Hooge, Zeelenberg, & Breugelmans, 2010; Frijda et al., 1989; Keltner & Buswell, 1997) and (b) seek to restore their positive image by performing well in another domain (de Hooge, Breugelmans, & Zeelenberg, 2008; de Hooge et al., 2010).

Second, initial setbacks may signal future challenges and make the prospect of an eventual failure loom large. Individuals with high performance expectations may be more likely than others to reduce persistence or quit a task in order to avoid becoming responsible for worse outcomes in the future. For example, Senate Minority Leader Harry Reid, who had led Senate Democrats for 10 years and pushed President Obama's ambitious agenda against Republican resistance, announced in 2015 that he would not seek reelection. Though he faced a difficult re-election bid and political landscape, Reid explained that his decision to leave politics was not motivated by concerns about winning re-election. Media sources, however, speculated that Reid's "retirement avoids him a humiliating defeat" (Meyers, 2015).

Third, quitting can also be a 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 future poor performance (Berglas & Jones, 1978; Greenberg, Pyszczynski, & Paisley, 1984; Kolditz & Arkin, 1982). In particular, the decision to reduce persistence after 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. These excuses offer an alternative account for poor performance, separate from low skill, especially in ambiguous situations where observers cannot easily judge why initial performance was poor. By reducing persistence and offering an excuse, individuals take a pre-emptive action that makes the link between their poor performance and their true ability ambiguous (Berglas & Jones, 1978). For example, students may attend parties, rather than study prior to a difficult exam; when they perform poorly, the link between their performance and true ability is unclear.

Practically speaking, executives frequently employ quitting as an image management strategy. For example, Chandini Portteus, the CEO of the Livestrong Foundation, was recruited from the Susan G. Komen Foundation to rebuild Livestrong's reputation after Lance Armstrong's doping scandal (Lindsay, 2015). Just nine months into her tenure at Livestrong, Ms. Portteus abruptly resigned, citing her need to focus on her family. Media reports speculated that her real concern was her inability to revive and lead the Livestrong Foundation (McCambridge, 2016). In this case, by pointing to a need to focus on family, Portteus offered an excuse for poor performance, and by quitting, avoided responsibility for the prospect of an even worse outcome at the Livestrong Foundation as it continued its slide.

Hypotheses

Although high performance expectations can benefit individuals in many ways (e.g., by enhancing motivation and performance; McNatt, 2000; McNatt & Judge, 2004), we argue that past research has overlooked the important possibility that high expectations can also *disadvantage* individuals by creating image management challenges. We theorize that compared to individuals with low performance expectations, those with high expectations will feel greater pressure to avoid poor performance; this is because poor performance would represent a greater violation of what is expected of them, causing more damage to their self- and public image. Poor initial performance may lead individuals with high expectations to feel more embarrassed and ashamed and thus reduce persistence as an image management tactic in order to (i) alleviate embarrassment and shame produced by underperforming, (ii) prevent a bigger failure from materializing, and (iii) provide an opportunity for offering an excuse for their poor initial performance. Formally, *Hypothesis 1: Individuals with high performance expectations will be less persistent than those with low performance expectations following weak initial performance.*

Hypothesis 2: Individuals with high performance expectations will experience more embarrassment and shame than those with low performance expectations following weak initial performance.

However, we do not expect to observe a difference in persistence between those with high and low performance expectations following strong initial performance. Strong initial performance does not elicit negative self-conscious emotions. Thus, we predict that the effects of performance expectations on both negative self-conscious emotions and persistence will be moderated by the quality of initial performance.

Hypothesis 3: The effect of performance expectations on persistence will be moderated by initial performance whereby people with high performance expectations will be less persistent than those with low performance expectations when initial performance is weak but not when initial performance is strong.

Hypothesis 4: The effect of performance expectations on embarrassment and shame will be moderated by initial performance whereby people with high performance expectations will experience more embarrassment and shame than those with low performance expectations when initial performance is weak but not when initial performance is strong.

OVERVIEW OF STUDIES

We present one field study (Study 1) and two laboratory studies (Studies 2 and 3) that investigate how people's performance expectations influence their persistence decisions and image concerns (measured through feelings of embarrassment and shame). We begin by examining a high-stakes, natural setting where expectations and the quality of initial performance vary substantially: professional men's tennis. We analyzed data on all available professional matches played between 1978 and 2011, and we used a regression discontinuity design to *causally* test whether being the favorite (versus the underdog) in a match increases the likelihood of quitting after falling behind in a match (Hypothesis 1) and whether this difference disappears when a player instead acquires an early lead (Hypothesis 3). In Study 2, we examined how people's performance expectations influence their decisions to persist in a trivia contest following poor initial performance, testing Hypothesis 1. The purpose of Study 2 was to conceptually replicate the key findings of Study 1 in a different environment to establish the generalizability of the phenomenon. Finally, in Study 3, we investigated the psychological underpinnings of the relationship between expectations and persistence in a multi-round competition. We examined how performance expectations and the quality of initial performance jointly influence concerns about one's self- and public image by measuring embarrassment and shame (Hypotheses 2 and 4).

STUDY 1: A FIELD STUDY OF PERFORMANCE EXPECTATIONS AND PERSISTENCE

To test our hypotheses about the effects of performance expectations on persistence (Hypotheses 1 and 3), we sought a competitive setting in which individual competitors had clear performance expectations (in this case established through world rankings) and persistence decisions were observable (in this case by observing quitting mid-task). We chose a domain with decades of rich data that was uniquely well-suited for testing our predictions: professional men's tennis. Failing to persist (by quitting) in this setting has no immediate economic advantages over losing; for professional tennis players, quitting and losing have identical implications for world rankings and tournament outcomes (The Association of Tennis Professionals, 2016). That is, compared to experiencing a loss, quitting yields identical ranking and tournament outcomes, but has potentially different impression management consequences because quitting implies that the player is injured. Injury is a potential excuse for disappointing performance. (Note that according to the Association of Tennis Professionals' (ATP) rules, the only allowable reason for quitting mid-match is injury).

Dataset

We compiled our dataset from the online data archive maintained by the ATP World Tour (www.atpworldtour.com). Our dataset includes 328,515 men's professional singles tennis matches played between 1973 and 2011 that recorded information critical to our research questions: players' world rankings and match outcomes. 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 (men's professional tennis matches nearly all consist of a best two-out-of-three set competition), but we do not know the scores of individual games or the order in which games were won or lost during a set. Critically, our data has information about whether either player in a match quit mid-match, which we use to measure a player's persistence, as well as the match score at the time when a player quit. Our dataset records 1,849 matches that did not actually start or ended before any score was recorded due to players withdrawing prior to a match or other rare reasons (e.g., suspensions).

Analysis Strategy

We analyze these data to test our prediction that holding all else equal (e.g., player quality/skill, age, etc.), higher performance expectations decrease persistence (i.e., increase the likelihood of quitting mid-match) after poor initial performance (Hypothesis 1), but not after strong initial performance (Hypothesis 3). In the context of tennis, players and observers have clear and consistent expectations for performance. These expectations are informed by players' ATP world rankings,[1] which are established by the Association of Tennis Professionals. In a tennis match, the player with a better ATP ranking is expected to win (i.e., has higher performance expectations) and is called the "favorite", whereas the player with an inferior ATP ranking is expected to lose (i.e., has lower performance expectations) and is called the "underdog."

Based on past research confirming that people use rankings to inform performance expectations, we expect that tennis players and observers have strong beliefs about the predictive power of rankings and favorite status. First, past research shows that players in a competition respond to small ranking differences even when the underlying skill differences they reflect are arguably meaningless or random. For example, Legge and Schmid (2014) show that merely reaching the podium (i.e., winning 3rd place) in a competition significantly increases professional skiers' performance in their next competition, compared with barely missing the podium (i.e., coming in 4th place). Though achieving 3rd versus 4th place reflects subtle, non-informative differences in race times, skiers respond to the label of being a prize winner. In addition, the observers of competitions are often overly-sensitive to small differences in rankings. They base their decisions on rankings even when (i) rankings provide no additional information beyond the underlying performance scores used to generate them (e.g., Luca, 2014; Pope, 2009) and (ii) changes in rankings are not accompanied by changes in objective quality measures (e.g., Luca & Smith, 2013). Further, people are sensitive to boundaries in rankings even when options barely

[1] We use ATP rankings rather than tournament seedings (where a "seed" is a status bestowed upon the highest-ranked subset of players in a tournament indicating their expected performance in the tournament) for a number of reasons. First, both players competing in a match are only seeded in 7.8% of all observations in our dataset. Thus, if we use seeding to determine which player was the favorite in a match, our sample size shrinks substantially. Second, for most tournaments, seeding directly reflects ATP rankings. Third, rankings tend to be the superior measure of which players are the best at a certain time, according to the ATP website (http://www.atpworldtour.com/Rankings/Rankings-FAQ.aspx).

within the boundary are extremely similar to options that that are barely outside the boundary (Isaac & Schindler, 2014). Altogether, this past research suggests that competitors and their observers both view small ranking differences as meaningful signals of differences in quality (or winning likelihood), even when this is inaccurate.

In this study, we examined (a) whether a tennis player who has initially performed poorly in a match (who thus faces a probable loss) is more likely to quit that match when he is classified as a favorite as opposed to an underdog (Hypothesis 1), and (b) whether this effect is eliminated when a player has initially performed well (and faces a probable victory; Hypothesis 3).

Identifying the impact of being a favorite rather than an underdog on a player's likelihood of quitting is not a simple task 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) or self-selection issues (e.g., a player with injuries 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 where the player who is the favorite to win a match is randomly assigned; of course, this never happens. To approximate random assignment to favorite status, we use a quasi-experimental sharp regression discontinuity (RD) design. The 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 involves assigning players to either favorite or underdog status based on a continuous assignment variable that reflects players' likelihood of winning a match and captures players' relative skills. The continuous assignment variable we rely on is a transformed *rank ratio* variable that we calculate for each player in each match, amounting to the target player's opponent's ranking divided by the target player's own ranking. We take the logarithm transformation of this *rank ratio* to derive a measure we refer 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 $-0.0100 = \log\left(\frac{100}{101}\right)$. Importantly, the *log rank ratio* is a smooth, linear predictor of a player's likelihood of winning a match (see Figure 1 Panel A).[2] In regressions predicting winning likelihood, *log rank ratio* is significant and positive ($\beta = 0.13$, $p < .0001$), and the squared term is not statistically significant ($p = 1.00$). The R-squared is higher with *log rank ratio* predicting the likelihood of winning (Adjusted R-squared = 0.0924) than it is with other intuitive measures, such as the difference between two players' rankings (Adjusted R-squared = 0.0654) or the ratio of two players' rankings (Adjusted R-squared = 0.0073). Neither of these alternative measures are well-behaved predictors (smooth and linear) of who will win a match (see Appendix A). Thus, *log rank ratio* is an excellent measure of a player's expected likelihood of winning a match and is superior to alternative, intuitive measures, such as differences or raw ratio of players' rankings. Therefore,

[2] This smooth linear relationship is important because a choppy relationship near the favorite/underdog threshold in winning likelihood could produce our main result in this study (a discontinuity in quitting) for reasons that are not purely psychological.

we use *log rank ratio* as the continuous assignment variable in our RD design and analyze the threshold at which players are tied in rank (i.e., *log rank ratio* = 0):[3] A *log rank ratio* greater than 0 indicates that the target player has a better (i.e. lower) ATP ranking than his opponent, and is classified as the favorite; a *log rank ratio* less than 0 indicates that the target player has a worse (i.e. higher) ATP ranking than his opponent, and is classified as the underdog.

Importantly, 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 players' likelihood of quitting. Players very near the rank equality threshold, but just on either side of the rank equality threshold are practically identical in terms of skill level. For example, imagine a match with players ranked 300 and 301. In our design, we assign the player ranked 300 "favorite" status and his opponent ranked 301 "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 differences like these as meaningful (e.g., Isaac & Schindler, 2014; Legge & Schmid, 2014; Luca, 2014; Luca & Smith, 2013; Pope, 2009). We expect these differences to inform judgments of outcome expectations. That is, we expect to find that trivial differences in ranking will significantly and discontinuously influence players' persistence. If we find this, 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 takes a value of 1 when a given player quits in a given match, and 0 otherwise. Quitting during a match is a stark measure

---

[3] Note, however, that our results are robust to using alternative continuous assignment measures such as the difference between two players' rankings and the raw ratio of two players' rankings (see Appendix C for these robustness tests).

of persistence. To capture whether a player has higher or lower performance expectations than his opponent, we create an indicator variable, *favorite*, which equals 1 if a player has 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 and causally affects a player's likelihood of quitting. Our dataset includes 628 matches in which both players have the same recorded rank. In these rare cases, we are unable to label either player the favorite or underdog in a match. In our results, we exclude these matches, but our results are meaningfully unchanged if we treat both players in these matches as favorites or both as underdogs.

To assess whether or not a player has performed poorly early on in a match, we use an objective measure of initial performance. In tennis, players who win the first set of a match have a high probability of winning the entire match, on average. Across all matches in our data set, players who win the first set, win the entire match 82% of the time. In our analyses, we use defeat in the first set as our binary measure of poor initial performance. We examine how this first set loss moderates the effect of favorite status on persistence. We create a dummy variable, *first-set loser*, to indicate the loser of the first set of a given match. For our regression models that include this variable as well as models that only include players who lost (or won) the first set, we exclude 4,218 matches in which players did not finish the first set.

We include a number of important control variables in all of our analyses. First, we include a fourth-degree polynomial of our *log rank ratio* variable to fully control for the continuous relationship between a player's skills relative to his opponent's and his likelihood of quitting. To choose the order of this polynomial, we followed Chen and Shapiro's (2007) method: we iteratively added higher-order terms to our regressions up to the point at which the next-highest term was no longer statistically significant in predicting quitting in any of our main regression models reported in Table 2. When estimating the moderating effect of losing the first set, we include interactions between the indicator for whether a player lost the first set and each term of the fourth-degree polynomial of log rank ratio (four terms total). Including the polynomial of log rank ratio and the interaction between each term of the fourth-degree polynomial and the indicator for losing the first set in our regressions ensures that the interaction between the indicator for losing the first set and the favorite indicator is not picking up a spurious effect of our key predictor—favorite status—on quitting (i.e. a continuous effect of *log rank ratio* rather than a discontinuous difference between favorites and underdogs).

We also control for the exact score in the first set (by including fixed effects for first set score: 0-6, 1-6, 2-6, etc.)[4] to compare players with the same outcomes in the first set of a match. In addition, we control 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 include 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 control 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 include controls for the round (1st, 2nd, 3rd, etc.) in which a match is played because later rounds in a tournament are likely to 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 include complete controls for the interactions between tour and round. Also, we control

[4] Note that we do not know the order in which games are won or lost during a match, only final set scores.

for court surface, which can affect injury (Girard, Eicher, Fourchet, Micallef, & Millet, 2007) and quitting rates (Breznik & Batagelj 2012). Finally, we control 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 include a second-order polynomial of player's age. Table 1 shows summary statistics for all matches in our sample that actually started. Appendix B provides additional information about control variables in our regressions. It is worth noting that our results regarding players' mid-match quitting decisions are robust to excluding all of the control variables detailed in this paragraph.

Regression Specifications

We use an ordinary least squares (OLS) model to predict quitting,[5] though our results are meaningfully unchanged when we instead rely on logistic regression models (and we report results from logit models in Appendix C as a robustness check). To address the concern that a binary outcome violates OLS assumptions of normal distributions and heteroscedasticity, we report bootstrapped standard errors based on 500 repetitions throughout the paper.

We begin by separately examining the effects of being a favorite on quitting rates among players who lost the first set and among players who won the first set. We only include one player from each match in these analyses: either the player who lost the first set or the player who won the first set. Formally, our model can be stated as follows, for a target player *i* in a given match *j*:

[5] We report the result from ordinary least squares 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). Further, we include interaction terms in our models, and interaction terms in nonlinear models can be subject to bias, as highlighted in Ai and Norton (2003). However, as shown in Appendix C, our results are nearly identical when we instead rely on logistic regressions.

$$(1) \quad quit_{ij} = \alpha + \beta_1{*}favorite_{ij} + \beta_2{*}\log(rank\_ratio_{ij}) + \beta_3{*}\log(rank\_ratio_{ij})^2 + \beta_4{*}\log(rank\_ratio_{ij})^3 + \beta_5{*}\log(rank\_ratio_{ij})^4 + \beta_6{*}rank_{ij} + \beta_7{*}\log(rank\_ratio_{ij}){*}rank_{ij} + \beta_8{*}\log(rank\_ratio_{ij})^2{*}rank_{ij} + \beta_9{*}\log(rank\_ratio_{ij})^3{*}rank_{ij} + \beta_{10}{*}\log(rank\_ratio_{ij})^4{*}rank_{ij} + \beta_{11}{*}opponent\_rank_{ij} + \beta_{12}{*}\log(rank\_ratio_{ij}){*}opponent\_rank_{ij} + \beta_{13}{*}\log(rank\_ratio_{ij})^2{*}opponent\_rank_{ij} + \beta_{14}{*}\log(rank\_ratio_{ij})^3{*}opponent\_rank_{ij} + \beta_{15}{*}\log(rank\_ratio_{ij})^4{*}opponent\_rank_{ij} + \beta_{16}{*}age_{ij} + \beta_{17}{*}age_{ij}^2 + \beta_{18}{*}prize\_money_i + \boldsymbol{\delta}{*}\boldsymbol{X}_{ij} + \boldsymbol{\theta}{*}\boldsymbol{Z}_j + \varepsilon_{ij}$$

where $quit_{ij}$ takes a value of 1 when target player *i* quits in match *j*, and 0 otherwise. Our primary predictor variable is the indicator for being a favorite (vs. underdog). We expect the coefficient on this variable to be significant and positive among players who lost the first set (exhibiting poor initial performance), but we do not expect favorite status to significantly predict quitting rates among players who won the first set (exhibiting strong initial performance). We include all of our control variables. In this model, $\boldsymbol{X}_{ij}$ is a vector of control variables representing the first-set score from the target player's perspective. For example, if the target player won 6 games and his opponent won 3 games in the first set, the first-set score would be coded as 6-3 for the target player. Further, $\boldsymbol{Z}_j$ 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.

Next, we examine whether the extent to which the discontinuous effect of being classified as a favorite (vs. underdog) on quitting is greater when players lost the first set than when players won the first set. This analysis requires 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. As a result, a match that ends with one player quitting corresponds to one observed quitting decision and one observed persistence decision. To deal with the dyadic, non-independent structure of our data, we relied on multilevel modeling (Kenny, Kashy & Cook, 2006) to estimate the effects of player and match characteristics on quitting. Formally, our model can be stated as follows, in which individual players (indexed by *i*) are nested within matches (indexed by *j*):

$$\begin{aligned}
(2) \quad quit_{ij} = \; & \alpha_j + \beta_1*favorite_{ij} + \beta_2*first_set_loser_{ij} + \beta_3* favorite_{ij}*first_set_loser_{ij} + \\
& \beta_4*\log(rank_ratio_{ij}) + \beta_5*\log(rank_ratio_{ij})^2 + \beta_6*\log(rank_ratio_{ij})^3 + \\
& \beta_7*\log(rank_ratio_{ij})^4 + \beta_8*\log(rank_ratio_{ij})*first_set_loser_{ij} + \\
& \beta_9*\log(rank_ratio_{ij})^2*first_set_loser_{ij} + \beta_{10}*\log(rank_ratio_{ij})^3*first_set_loser_{ij} \\
& + \beta_{11}*\log(rank_ratio_{ij})^4*first_set_loser_{ij} + \beta_{12}*rank_{ij} + \beta_{13}*\log(rank_ratio_{ij}) \\
& *rank_{ij} + \beta_{14}*\log(rank_ratio_{ij})^2*rank_{ij} + \beta_{15}*\log(rank_ratio_{ij})^3*rank_{ij} + \\
& \beta_{16}*\log(rank_ratio_{ij})^4*rank_{ij} + \beta_{17}*opponent_rank_{ij} + \beta_{18}*\log(rank_ratio_{ij}) \\
& *opponent_rank_{ij} + \beta_{19}*\log(rank_ratio_{ij})^2* opponent_rank_{ij} + \\
& \beta_{20}*\log(rank_ratio_{ij})^3*opponent_rank_{ij} + \\
& \beta_{21}*\log(rank_ratio_{ij})^4*opponent_rank_{ij} + \beta_{22}*age_{ij} + \beta_{23}*age_{ij}^2 + \boldsymbol{\delta}_{\mathbf{ij}}*\boldsymbol{X}_{ij} + \varepsilon_{ij}
\end{aligned}$$

$$(3) \quad \alpha_j = \gamma_0 + \gamma_1 *prize\_money_j + \boldsymbol{\theta}_{\mathbf{j}}*\boldsymbol{Z}_j + u_j$$

Here as before, $quit_{ij}$ takes a value of 1 when target player *i* quits in match *j*, and 0 otherwise. Equation (3) estimates intercepts for matches based on match-level data, and equation (2) estimates the outcome of interest (quitting) based on match-level and individual-level data. The level-1 predictor variable of interest is the interaction term between the favorite indicator and an indicator for losing the first set, which we expect to have a positive and significant coefficient. In addition to controlling for other variables in (1) that vary between players within the same match *j*, we add level 1 predictor variables for the interactions between *first_set_loser* and each term of the fourth-order polynomial of log rank ratio. Level 2 predictor variables include prize money and other control variables that vary at the match level (denoted by a vector $\boldsymbol{Z}_j$) including:

tournament round, tour, round-tour interaction, court surface, and year. Finally, ε_{ij} and u_j are error terms at level 1 and level 2, respectively.

Results

The average rate of quitting mid-match is 1.39% (or 9,132 matches) in our data set. In Figure 2 Panel A, we show quitting rates for the 50% of matches in our dataset where competitors' rankings were the closest. We group observations into bins based on players' *log rank ratio*. Each dot represents the probability that a player quit in the middle of a match (vertical axis) according to the designated bin for the player's *log rank ratio* (horizontal axis). Panel A of Figure 2 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, $\chi^2 = 3.86$, $p < .05$.[6] The increase in the likelihood of quitting by 18.60% at this underdog-favorite threshold is notably larger than the increase we observe at other rank-ratio thresholds. However, as Panel B of Figure 2 shows, those slight underdogs and slight favorites exhibit virtually identical quitting rates after exhibiting strong initial performance and winning

[6] 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 see the same pattern whereby favorites retire at a significantly higher rate (2.09%) than underdogs (1.62%), χ^2 (1) = 12.91, $p < .001$.

the first set (0.53% for slight favorites and 0.52% for slight underdogs, $p = 0.84$), consistent with Hypothesis 3.

We next report a series of regression models that take into account relevant control variables to test Hypotheses 1 and 3. Model 1 in Table 2 follows the regression specification (1) and focuses on players who exhibited weak initial performance and lost the first set. The positive and significant coefficient estimate on *favorite* ($\beta = .0037$, $p < .001$) indicates that players who lost the first set have markedly different patterns of quitting behavior right at the threshold of equal rank, in support of 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). However, Model 2 in Table 2 suggests that favorites are no more likely to quit than underdogs if they exhibited strong initial performance, winning the first set ($\beta = .0002$, $p = .72$), consistent with Hypothesis 3.

Next, we test whether the effect of being a favorite on quitting significantly differs between players who had poor initial performance (losing the first set) versus strong initial performance (winning the first set). Model 3 in Table 2 applies the multi-level regression specification in equations (2) and (3) and reveals that the interaction between exhibiting poor initial performance (losing the first set) and being the match favorite is significant and positive ($\beta = 0.0054$, $p < 0.0001$).[7] This indicates that the discontinuous jump in quitting rates at the

[7] 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 standard error of 0.0008, $p < 0.0001$). This estimated effect size, within the margin of error, is similar to the effect estimated in Model 1 (i.e., 0.0037) 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 Models 2 (coefficient on *favorite* = 0.0002, bootstrapped standard error = 0.0005) and 3 (coefficient on *favorite* = -0.0005, bootstrapped standard error = 0.0005), within the margin of error.

underdog-favorite threshold is significantly stronger for favorites who lose the first set than for favorites who won the first set. These results support Hypothesis 3.

Testing the validity of the RD design. Thus far, we have relied on a RD design to show that being the favorite (rather than the underdog) in a competition significantly increases a player's likelihood of quitting when the player exhibits poor initial performance but not when the player initially performs well. To claim that we have *causal* evidence of this fact, we need to ensure that we satisfy the RD design's critical assumption that around the equal rank threshold, being assigned to favorite versus underdog status is essentially random. To test this assumption, we conduct two standard tests (Imbens & Lemieux 2008).

The first RD test involves contrasting players immediately above and below the equal rank threshold. To satisfy this test, we need to show that player characteristics in the two groups are very similar. In our setting, we need to compare characteristics, separate from the *log rank ratio*. Apart from players' rankings, which we used to calculate *log rank ratio*, the only ex-ante characteristic that we observe that varies across players in the same match is their age.[8] We do observe that slight underdogs (*log rank ratio* between -0.10 and 0) are 0.11 years younger (M = 23.78) than slight favorites (*log rank ratio* between 0 and 0.10; M = 23.89, p = .0003). This difference is small, and we postulate that this one-month difference cannot account for our results. Still, we include age and age squared as controls in our regression analyses.

We also consider whether a discontinuity in relative skill levels exists at the favorite-underdog threshold and successfully rule out this possibility. Since we do not directly observe relative skill levels, we use a player's likelihood of winning a match as a proxy. We rely on

[8] There is no need to examine characteristics for which both players in a match have the same value (e.g., tournament round, prize money, tournament type) because for those characteristics, a favorite must have the same value as the underdog in the same match by definition.

identical regression specifications to those presented in Models 1-3 of Table 2 to examine the relationship between skill and favorite/underdog status, except that the new regressions predict whether a player wins a match instead of whether he quits (Models 4-6 in Table 2). In these regressions, *favorite* is not a significant predictor of the likelihood of winning for players who lost the first set (Table 2, Model 4), for players who won the first set (Table 2, Model 5), or in our interaction model (Table 2, Model 6). Also, there is not a significant interaction between *favorite* and the indicator for losing the first set in Model 6, Table 2. Thus, we can confirm that there are not significant and discontinuous differences in the odds of winning (our proxy for skill) between slight favorites and slight underdogs.

The second standard RD test of the randomness assumption evaluates whether there is a discontinuity in the distribution of *log rank ratio* at the rank equality threshold. Such a discontinuity would suggest nonrandom assignment of favorites and underdogs to their positions around the threshold of interest. Following Imbens and Lemieux (2008), we group player-match observations into bins based on their *log rank ratio* and plot the number of observations in each bin. We focus on the 50% of observations that are closest to the equal rank threshold (i.e., *log rank ratio* falls between -1 and 1). Figure 1 Panel B shows that as the absolute value of a player-match's *log rank ratio* moves from 1 to 0, the number of observations in a bin first increases gradually and then decreases for the two bins right next to the favorite-underdog threshold (e.g., bins -0.1 to 0 and 0 to 0.1). Importantly, the change in the number of observations is smooth across the rank equality threshold, and we observe no discontinuity in the distribution of *log rank ratio* when comparing favorites whose *log rank ratio* is slightly above 0 with underdogs whose *log rank ratio* is slightly below 0. In addition, by construction, we observe symmetry around the underdog-favorite threshold (where *log rank ratio* = 0).

The discontinuity test above is intended to address the concern that players self-select from one side of the favorite-underdog threshold to the other, which, were it to occur, would invalidate the randomness assumption underlying RD designs. 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*). It is, in principle, possible that a discontinuity in pre-match withdrawals between underdogs and favorites could explain the discontinuity we detect in mid-match quitting. The average rate of pre-match withdrawals is 0.19% in our data set. To test the (unlikely) possibility that a discontinuity in pre-match withdrawals could be driving our effects, we ran a regression similar to the multi-level mixed effects model we detail in equations (2) and (3) to predict a player's choice to withdraw prior to a match as the binary dependent variable. The Level-1 and Level-2 predictor variables are similar to those in equations (2) and (3) except that we exclude the *first_set_loser* indicator, its interaction with the favorite indicator and the polynomials of log rank ratio, as well as control variables representing first-set scores, because matches with a pre-match withdrawal did not start and thus did not have information about a player's performance in the first set. Model 7 in Table 2 shows that *favorite* is not a significant predictor of pre-match withdrawals, indicating that there is no evidence that favorites and underdogs near the rank equality threshold withdraw prior to matches at a different rate. Thus, we can rule out selection effects as a possible explanation for our findings.

Ruling out discontinuities at other thresholds as an explanation for quitting behavior. We next consider whether or not quitting is so unpredictable that statistical tests spuriously reveal 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 detect 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, Dahl, and Nielson (2013) and reran the regression specified in Model 1 of Table 2 twenty-one 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, with X 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 players lose the first set. Figure 3 presents the estimated coefficients on the modified favorite indicator for each of these regressions twenty-one 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.

Robustness tests. To ensure the robustness of our results, we conducted several robustness checks and report results regarding players' mid-match quitting decisions in

Appendix C. First, our results remain robust when we rely on logistic regression models instead of OLS regressions.[9] Second, our results are robust if we adjust the operationalization of key 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, or (c) treating both players in a match as underdogs if they have identical rankings. Third, our results remain robust if we remove extremely high or low *log rank ratio*s by dropping all observations with a *log rank ratio* above the 97.5th percentile or below the 2.5th percentile. In addition, throughout our analyses we rely on players' *log rank ratios* to capture their relative skills, but our results are robust to using differences in ranks or players' raw rank ratio instead of their log rank ratio (though these alternative metrics are inferior given that they are not linear predictors of match victory, as we discussed in Analysis Strategy Section).

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) present 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 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 discontinuity threshold, as recommended by Gelman and Imbens (2014).

[9] For regression models that only include one player from a match, we obtain robust results from logistic models. However, for regression models that include both players in a match and thus rely on multi-level mixed-effects models, we are unable to run multi-level mixed-effects logistic regressions due to the computational constraints required by such models because our level-2 variable has a large number of groups (i.e., 323,683 unique matches). Thus, we report results from simple logistics regressions that do not incorporate multi-level mixed effects.

Specifically, to conduct this robustness test, we discard observations with a *log rank ratio* value that is greater than a selected bandwidth away from our threshold of zero 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 estimate a linear function on remaining observations whose *log rank ratio* is near zero (Hahn, Todd, &Van Der Klaauw, 2001). We follow Imbens and Kalyanaraman (2009, 2012) to derive the "optimal bandwidth" around the rank equality threshold for this test. Hypotheses 1 and 3 remain supported when we switch to this alternative empirical approach and Appendix D provides 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 begin to lose, favorites who begin to lose are more likely to quit.

The effect we detect is substantial—being a slight favorite (versus a slight underdog) increases the likelihood of quitting among first-set losers by 26.62% relative to the average quitting rate observed in our data. Further, the effect is robust to a number of alternative specifications and robustness tests. Importantly, our regression discontinuity design does *not* compare average favorites with average underdogs. Rather, it compares favorites and underdogs who are *just* above or below the underdog/favorite threshold. This approach allows us to rule out the possibility that underlying differences between underdogs and favorites (e.g., in endorsements, riches, ego, number of fans, skills, etc.) could account for our finding, because underdogs and favorites who lie just above versus below the threshold are virtually identical across these characteristics. We can also rule out selection effects for our findings since tennis players cannot directly influence the opponent they face and we observe 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, and we find support for Hypotheses 1 and 3. Next, we present a laboratory experiment to test Hypothesis 1 in a very different setting.

STUDY 2: PERFORMANCE EXPECTATIONS AND PERSISTENCE IN THE LABORATORY

We extend our investigation to a very different setting. In this experiment, we manipulated individuals' performance expectations and measured their persistence decisions in a trivia contest. We focus on replicating our finding in Study 1 that following poor initial performance, the persistence decisions people make differ as a function of their performance expectations such that those with higher expectations will be less persistent than those with low expectations.

Participants

We conducted this study at a large Midwestern university in the United States. Participants received course credit for participating in the experiment and earned additional money depending on their performance. We decided in advance to recruit as many undergraduate student participants as possible in five days. The final sample included 201 participants (104 females, average age of 19 years).[10] The number of participants who attended each experimental session ranged from 5 to 12.

Experimental Procedure

[10] Among the 201 participants who started the experiment, three participants experienced technical errors and were dismissed from the experiment. Demographics information was not collected from those participants. Since our behavioral measure of interest in this experiment was persistence (explained later), these three participants were included in our analysis to ensure an intention-to-treat approach. Our results are robust if we exclude these three participants from our analysis (see Appendix E).

We adapted experimental procedures from prior studies of tournaments (e.g., Bull, Schotter, & Weigelt, 1987; Lim, Ahearne, & Ham, 2009), drawing heavily on studies that sought to create a social environment "close to most tournaments in the real world" (Chen, Ham & Lim, 2011, p. 869; Lim, 2010). At the start of each experimental session, the experimenter asked participants to stand in a circle and introduce themselves to the other participants by stating their name and academic major (Chen et al., 2011; Lim, 2010). Then, the experimenter handed out and read aloud instructions about a trivia contest.

We informed participants that they would be randomly assigned to compete in a trivia contest including questions of either "middle-school difficulty" or "expert difficulty." Participants learned that those in the *middle-school difficulty* condition would be presented with trivia questions that most people are exposed to by the end of middle school, and that good performers answer 70% of the questions correctly. Participants also learned that those in the *expert difficulty* condition would be presented with trivia questions that most people cannot answer, and that good performers answer 20% of these questions correctly.[11] Thus, by assigning participants to the middle-school difficulty or expert difficulty condition, we assigned them to have relatively high or low performance expectations, respectively. 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 Appendix F for the questions used in this study.

[11] In a pretest (N = 213) where participants only read the descriptions of the two conditions without information about good performers, participants assigned to the middle-school difficulty condition on average predicted that they would answer 70% of the questions correctly, whereas participants on average answered 20% of questions correctly in both conditions. Thus, we used 70% and 20%, respectively, as part of our performance manipulations.

Participants learned that they would have up to 30 seconds to select the correct answer to each trivia question. We told participants that the first 21 trivia questions would cover a broad array of topics, and that after the 21st question and after every question following that, they would be allowed to choose whether to continue answering trivia questions covering the same topics sampled thus far or switch to trivia questions on new topics. We told participants that regardless of whether or not they switched topics, the difficulty 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 lab session ended. 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 embarrassing) performance from being posted.

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 contest; 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 submitted after they switched. Thus, poor performing participants faced a persistence decision. 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 and answering any questions, she flipped a coin to determine participants' conditions. We asked participants in the middle-school difficulty condition to enter the computer lab first and sit at a computer terminal labeled "Trivia Challenge: Middle-School Difficulty." Next, the remaining participants entered the computer lab and sat at a computer terminal labeled "Trivia Challenge: Expert Difficulty."

Before beginning the trivia contest, participants predicted the percentage of questions they would correctly answer. Then they started the contest. After each question, we gave participants feedback so they knew whether or not their answer was correct and what percentage of questions they had answered correctly to that point. After participants received feedback on the 21st question, we presented them with their first opportunity to quit answering questions on the original set of topics and switch to questions about a new set of topics (while earning a lower piece-rate payment per correct question). Participants who continued with the original topic set were offered this switching opportunity after every question until the end of the contest (which included 45 questions). Notably, participants were not informed about how many questions were included in the contest or what new trivia topics they would encounter if they chose to switch topics.

Our primary 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 between 21 (for participants who switched topics at the first opportunity) and 45 (for participants who never switched topics). To address the concern that this measure is right-censored at 45, we also examined a binary outcome variable: whether participants switched topics when they were offered the first opportunity to do so after the 21st question.[12]

---

[12] For the three participants who were dismissed from the experiment early due to technical errors with our trivia program, the persistence measure was set equal to the number of questions they answered before encountering an error. This turned out to give us a conservative test of our Hypothesis 1 because the two participants in the expert difficulty condition who encountered technical issues with the quiz encountered these problems earlier than the participant in the middle-school difficult condition.

After finishing the trivia contest, participants answered a series of demographic 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.

Results

Manipulation check. Participants in the middle-school difficulty condition predicted that they would correctly answer 66.90% of the trivia questions and participants in the expert difficulty condition predicted that they would correctly answer 25.45% of the trivia questions, $t(199) = 18.57$, $p < .0001$. Thus, assigning participants to the middle-school vs. expert difficulty trivia questions successfully manipulated their performance expectations.

Persistence. Across the first 21 questions, participants in the middle-school difficulty and expert difficulty conditions both performed quite poorly (answering 25.67% of questions correctly in the middle-school difficulty condition versus 29.61% in the expert difficulty condition, $p = 0.12$). Following Hypothesis 1, we predicted that when faced with poor initial performance, people with high performance expectations (those in the middle-school difficulty condition) would be less persistent than those with low performance expectations (those in the expert difficulty condition). Indeed, as shown in Figure 4, we found that participants in the middle-school difficulty condition—who began the contest with higher performance expectations—switched topics sooner ($M_{persistence} = 24.08$, $SD = 7.18$) than did participants in the expert difficulty condition ($M_{persistence} = 27.71$, $SD = 10.07$), $t(199) = 2.94$, $p = .004$. Further, participants in the middle-school difficulty condition were significantly more likely to switch when they were offered the first opportunity to do so (64.00%) than were participants in the expert difficulty condition (41.58%), $\chi^2(1) = 10.13$, $p = .001$.

Discussion

In this experiment, we measured persistence by assessing how long participants in a trivia contest continued to answer questions on the same topics before electing to switch to questions on different (and less well-remunerated) topics. Our design mimicked common situations in organizations where employees choose to either persist on a current project or switch to a different one. This study builds on the results of our field study of men's professional tennis players by providing additional evidence in a distinct setting. When individuals experience poor initial performance, those with high performance expectations are less persistent than those with low performance expectations (consistent with Hypothesis 1).

In our next study, we extend our investigation with a second laboratory experiment, and explore the underlying mechanism that links performance expectations, initial performance, and persistence. Specifically, we examine how performance expectations and initial performance affect two emotions associated with image concerns: embarrassment and shame (Hypotheses 2 and 4).

STUDY 3: PERFORMANCE EXPECTATIONS, EMBARRASSMENT, AND SHAME

This experiment builds on Studies 1 and 2 by delving into the underlying mechanism. In this study, we assess the emotions triggered by poor performance by competitors with high and low performance expectations.

In our experiment, we conducted a multi-round competition. We manipulated participants' performance expectations, and we orthogonally manipulated the quality of participants' initial performance by varying whether or not they won the first round of the multi-round competition. That is, we used a 2 (Expectations: High vs. Low) x 2 (First round performance: Win vs. Loss) between-subjects design.

With this design, we can test Hypotheses 2 and 4. To test Hypothesis 2, we assess whether or not participants with high performance expectations (favorites) who lost in the first round feel more embarrassed and ashamed than participants with low performance expectations (underdogs) who lost in the first round. According to Hypothesis 4, we do not expect participants with high and low expectations who won the first-round to experience differences in these emotions.

Participants

We conducted this study at a large Northeastern university in the United States. Participants received $10 for participating in the experiment and earned additional money depending on their performance. We decided in advance to recruit as many undergraduate student participants as we could in three days. Our final sample included 156 participants who completed the study (101 females, average age of 20 years).[13] The number of participants who attended each experimental session ranged from 6 to 18.

Experimental Procedure

As was the case in Study 2, we started each session by acquainting participants with each other. The experimenter asked participants to stand in a circle and introduce themselves by stating their name, academic major, and an interesting event that they had experienced at some point in their life (Chen et al., 2011; Lim, 2010). Then, the experimenter handed out and read the study instructions aloud.

We informed participants that they would compete with another participant in a three-round "math contest" and that they would work on a series of addition problems in each round.

[13] Among the 160 undergraduates who began the study, four participants did not complete the contest because of technical errors that arose during the study. They did not respond to our dependent measure and thus could not be included in our analysis.

We told participants that for each addition problem, their task was to calculate the sum of five, two-digit numbers without using a calculator. Participants would win the contest and be paid \$0.20 for each correct answer they submitted across the three rounds if they won at least two out of three rounds. Otherwise, they would lose the contest and be paid \$0.10 for each correct answer they submitted. Next, we informed participants that one contestant in each pair would be randomly selected to be the "frontrunner" and would have an extra 30 seconds to solve addition problems in each round. Frontrunners were given 3 minutes and 30 seconds per round, while the other contestants were given 3 minutes per round. Our manipulation of favorite status was adapted from Chen et al., (2011) who created "frontrunners" in a contest by offering some individuals an advantage over other contestants.

Next the experimenter had participants draw a card with a pair number (ranging from 1 to X, with X being the number of pairs of participants in a session) and a letter (either A or B) written on it. Participants who drew the same pair number (e.g., 3A and 3B) were matched as competitors. In front of all participants, the experimenter flipped a coin to determine whether participants with the letter A or B would be the frontrunner in their pair. We asked the frontrunners (hereafter, the "favorites") to wear a sticker with the word "Frontrunner" on it, write down their first name in the "Frontrunner" column of a table on a large blackboard, and sit at a computer terminal labeled "Frontrunner." Similarly, we asked participants assigned to compete with a frontrunner (hereafter, the "underdogs") to write down their first name in the "Opponent" column on the blackboard and to sit at a computer terminal labeled "Opponent." At this point, the blackboard showed one row for each pair of participants and displayed their pair number, the first name of the favorite, the first name of the underdog, and three blank columns labeled "Round 1," "Round 2," and "Round 3."

Both favorites and underdogs perceived that favorites had a substantial advantage in the contest.[14] However, unbeknownst to participants, we exposed participants to one of two types of addition problems that differed substantially in their degree of difficulty.[15] We randomly varied whether each participant answered easier or harder questions than his opponent to manipulate whether a participant would exhibit poor or strong initial performance (in the first round of the competition). In pair numbers 1, 2, 3, 5, 7, and 9, the favorites answered easier addition questions, and the underdogs answered harder addition questions throughout the contest. We expected the favorites in these pairs to win every round. However, in pairs 4, 6, and 8, the favorites answered harder questions, and the underdogs answered easier questions for the whole contest. Despite the fact that favorites had more time in those pairs, we expected the favorites to lose every round.[16] In summary, our manipulations created four treatments: favorites who answered easy questions and were extremely likely to win, underdogs who answered difficult questions and were extremely likely to lose, favorites who answered difficult questions and were extremely likely to lose, and underdogs who answered easy questions and were extremely likely to win. Thus, participants all expected the favorites to win; however, we randomly assigned some favorites to lose.

After participants completed Round 1, the experimenter recorded everyone's score and announced the outcome for each pair aloud in several steps. First, the experimenter reminded

[14] In a pretest, participants (N = 86) from the same subject population as this experiment were assigned to be favorites or underdogs. Pretest participants— regardless of their favorite versus underdog status—predicted that favorites on average should win the contest 80% of the time, a chance that is significantly higher than 50% ($p <$ .0001). This confirmed the effectiveness of our manipulation of performance expectations.

[15] The last digit of the two-digit numbers used in easier questions was either 0 or 5 (e.g., 55+25+30+45+80). The last digit of the two-digit numbers used in harder questions could be any number between 0 and 9 (e.g., 54+28+37+43+86).

[16] Indeed, in this experiment, all 52 favorites who were supposed to win by design won the first round, and 25 out of 26 favorites who were supposed to lose by design lost the first round. We report results based on an intent-to-treat analysis such that our independent variable was whether a participant was assigned to a condition with the goal of ensuring they would win or lose Round 1 rather than whether the participant actually won or lost Round 1.

participants that the frontrunners had 30 seconds more than their opponents and were the expected winners in each pair. Then, the experimenter mentioned that there were some surprising results, and added two columns named “Expected outcome” and “Upset,” respectively, within the column named “Round 1” on the blackboard. Next, for each pair, the experimenter announced who was *expected* to win Round 1 and who *actually* won Round 1. If the favorite in a pair won Round 1, the experimenter indicated on the blackboard that this result was an “expected outcome” for that pair; if the favorite in a pair lost Round 1, the experimenter indicated on the blackboard that this result was an “upset” for that pair. We used the term “upset” because the media often uses this term to refer to the unexpected loss of a favorite in a competition. The experimenter also asked participants to stand up when the results for their pair were announced and asked winners to raise their hands when their names were called.

After the experimenter announced the results of Round 1, participants proceeded to Round 2 and then Round 3. After each round, the experimenter simply announced which participant in each pair had won the round. After completing all three rounds, we asked participants to recall how they felt after they were informed of their results in Round 1. They rated the extent to which they felt embarrassed, self-conscious, ashamed, or disgraced on a seven-point Likert scale (1 = not at all; 7 = very much). Based on the Differential Emotion Scale modified by Mosher and White (1981), we used the emotion labels “embarrassed” and “self-conscious” to measure embarrassment—negative feelings associated with violating social expectations and shaking one’s public image; and we used the emotions labeled “ashamed” and “disgraced” to measure shame—negative feelings associated with violating personal expectations and shaking one’s self-image (Tangney et al., 1996).

Results

We hypothesized that favorites would feel more embarrassed and ashamed after an initial loss than would underdogs (Hypothesis 2), but we did not expect to find this difference after an initial success (Hypothesis 4). To test these hypotheses, we used a 2X2 ANOVA to predict embarrassment and shame after Round 1 as a function of two independent variables: (i) whether a participant was the favorite or underdog in a pair, and (ii) whether the participant answered the easier addition questions (i.e., was randomly assigned to experience strong initial performance) or answered the harder addition questions (i.e., was randomly assigned to experience poor initial performance).

For both embarrassment and shame, there was a significant interaction between a participant's status as a favorite or an underdog and the participant's randomly assigned performance in Round 1 (embarrassment: $F(1, 152) = 8.32$, $p = .005$; shame: $F(1, 152) = 8.80$, $p = .004$). The planned contrast analysis shows that after being randomly assigned to experience poor initial performance, favorites felt more embarrassed ($M = 4.38$, $SD = 1.75$) than underdogs ($M = 2.63$, $SD = 1.49$), $F(1, 152) = 23.86$, $p < .0001$; favorites also felt more ashamed ($M = 3.31$, $SD = 2.26$) than underdogs ($M = 1.95$, $SD = 1.42$), $F(1, 152) = 18.93$, $p < .0001$. These results support Hypothesis 2. However, confirming Hypothesis 4, after being randomly assigned to experience strong initial performance, favorites and underdogs did not differ significantly in feelings of embarrassment ($M_{favorites} = 2.42$, $SD = 1.51$ vs. $M_{underdogs} = 2.13$, $SD = 1.15$, $F(1, 152) = 0.65$, $p = .42$) or shame ($M_{favorites} = 1.22$, $SD = 0.55$ vs. $M_{underdogs} = 1.17$, $SD = 0.62$, $F(1, 152) = 0.02$, $p = .88$). Figure 5 depicts average feelings of embarrassment and shame for each of the experimental conditions.

Discussion

In this experiment, we assigned participants to either high or low performance expectations and to either win or lose the first round of competition. Consistent with Hypothesis 2, we found that after an initial loss, favorites felt more embarrassed and more ashamed than underdogs, suggesting that favorites were more concerned about their self-image and public image than underdogs following poor performance. However, consistent with Hypothesis 4, after an initial win, favorites and underdogs did not differ in feelings of embarrassment or shame. These findings offer insight into the psychological underpinnings of the effects we document in Studies 1 and 2. It feels bad to lose, but the experience of embarrassment and shame following a loss is especially intense when it violates the expectation of winning.

GENERAL DISCUSSION

Across one field study and two laboratory experiments, we examine how performance expectations influence persistence decisions and the experience of embarrassment and shame (emotions associated with self-image and public image concerns). In Study 1, we analyzed over 30 years of archival data from professional men's tennis matches. We find that competitors who expect to win are more likely to quit than competitors who expect to lose following poor initial performance. In Study 2, we experimentally manipulated performance expectations in the laboratory and found additional evidence in a different setting that high performance expectations reduce persistence when initial performance is poor. In Study 3, we orthogonally manipulated performance expectations and the quality of initial performance and found that high expectations elevate feelings of embarrassment and shame following poor initial performance, but not following strong initial performance. The combination of field and laboratory data across our three studies provides strong support for both the internal and external validity of our findings. When the going gets tough, individuals with high performance expectations are more likely to quit than those with low performance expectations.

Theoretical and Practical Implications

Our research makes several important contributions to management theory. First, throughout our lives—at work and at home—we make decisions about when to quit and when to persist. In spite of the enormous importance these decisions can have, there are significant gaps in our understanding of when and why individuals persist. We identify image management considerations as an important antecedent of quitting that may cause individuals with high expectations to be less persistent.

Second, in contrast to the extant literature that has overwhelmingly highlighted the advantages of having high performance expectations (Bandura, 1982, 1997; Eden, 1990, 1992; McNatt, 2000; McNatt & Judge, 2004; Merton, 1948; Rosenthal & Jacobson, 1968a, 1968b), we both theorize and demonstrate that individuals with high expectations face a previously unappreciated and significant obstacle to success. Following poor initial performance, high expectations amplify image management concerns, elicit feelings of embarrassment and shame, and cause people to quit more often than they would had their expectations been lower, even when quitting entails forgoing opportunities to earn substantial rewards (e.g., substantial prize money in the case of Study 1 and Study 2). Our findings add to the small but growing literature that has begun to recognize drawbacks to holding high performance expectations (e.g., Baumeister, Hamilton, & Tice, 1985; Mishina, Dykes, Block, & Pollock, 2010).

Our findings have a number of important implications for managers. Within organizations, managers and employees constantly face decisions to either persist or redirect their effort. Our findings reveal that it is critical for managers to understand how performance expectations shape these persistence decisions. These decisions manifest themselves in many different ways. For example, executives may step down from a leadership position, quit a competition, terminate an enterprise, stop mentoring a junior employee, or reduce their involvement in a project. Similarly, employees may reduce persistence in a task by reallocating effort from one task to another. This is particularly relevant to the contemporary multi-task work environment, and reduced persistence via effort re-allocation has been observed when employees feel discouraged by their performance on one task and want to improve their image by excelling on another task (Barankay, 2012). Regardless of its specific form, failures to persist in the workplace can be costly to both individuals and organizations. Importantly, 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.

Our findings shed light on how managers should manage subordinates' performance expectations. Past research suggests that mangers can strengthen employees' confidence and boost their motivation and job performance by communicating positive expectations (see McNatt, 2000 for a meta-analysis). However, our findings reveal that conveying positive expectations may come at a cost. If employees with high expectations experience initial setbacks, they may feel embarrassed and ashamed, and may reduce their effort. To reduce the likelihood of premature quitting and inducing negative emotions, managers may want to explore strategies that help employees with high expectations overcome image management concerns, such as fostering a culture where people feel that they can learn from failures or emphasizing processes more than outcomes. Similarly, our findings suggest that managers should allocate their scarce resource of attention to employees who experience early setbacks. Image concerns may prompt these employees to conceal relevant information and quit prematurely.

Our findings also identify a harmful consequence of ranking systems. Greater access to data and analytical tools has facilitated increased reliance on ranking systems in many industries and organizations (Mills & Mills, 2014). Ranking systems communicate expectations and may therefore impose image management concerns on those with high expectations. Our findings suggest that ranking systems should be used with caution and add to an ongoing discussion about the advantages and disadvantages of ranking employees (Barankay, 2012; Charness, Masclet, & Villeval, 2014) and of making rankings transparent (Bernstein & Li, 2016; Song, Tucker, Murrell, & Vinson, 2015).

Limitations and Future Directions

We employed a mixed methodological approach in this paper, combining analyses of laboratory and archival data to establish the robustness of our findings and their relevance to management practice. One concern might be that our field data is drawn from a professional sports context, rather than a more traditional organizational environment. In fact, many management scholars have successfully used sports data to study organizational phenomena (e.g., Berman, Down, & Hill, 2002; Cotton, Shen, & Livne-Tarandach, 2011; Kilduff, Elfenbein, & Staw, 2010; Larrick, Timmerman, Carton, & Abrevaya, 2011; Pope & Schweitzer 2011). One key advantage of sports data is that individual performance metrics (and quitting in our context) are readily observable. In addition, the sports context we studied represents a large, complex “organization” where many stakeholders and interest groups (e.g., players, coaches, sponsors, spectators, and the media) are heavily involved. Tennis in the United States is estimated to be a $5.57 billion business (Tennis Industry Association, 2013). Further, there are many similarities between the organization of tennis and corporations. For example, tennis players and CEOs are rewarded based on their performance, exposed to public scrutiny, subject to image management concerns, and accountable to stakeholders (e.g., sponsors versus venture capital investors). Still, future research that explores the relationship between expectations and persistence in corporate contexts would be valuable.

In the context of men's professional tennis, we note that 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 are discontinuously more likely to quit and claim an injury than slight underdogs after losing the first set, even though these competitors are identical (e.g., in terms of their likelihood of getting injured and the seriousness of their injuries) beyond their designation as "favorite" and "underdog." Thus, we have suggestive evidence that some of the injuries that favorites claim are, 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 image management strategy. However, there may be other mechanisms that contribute to our findings about tennis players' quitting decisions. For example, 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 have been injured, compared with underdogs. Such differences in attributions may cause favorites to quit more frequently than underdogs after losing the first set. Our tennis study cannot disentangle our hypothesized image management motive (i.e., players quitting and falsely claiming injuries to manage their images) from the aforementioned attribution-based explanations. However, in Study 3 we provide evidence that high versus low expectations are associated with different image concerns, consistent with our hypothesized mechanism.

One question that comes out of this research is whether the expectation-persistence relationship documented here represents a drawback to having high expectations or a benefit of having low expectations. One interpretation of our results is that individuals with low expectations have little to lose from a failure and much to gain from a success and thus are willing to stick with their current course of action, even when their odds of success are low. Such an interpretation of our findings suggests that by being persistent, individuals with low expectations not only retain the possibility of advancing and eventually succeeding but also may enjoy the long-term benefits associated with continued practice. In this work, we have focused on documenting the novel, negative relationship between expectations and persistence. Future research that differentiates between the psychology of individuals with high expectations and low expectations will be valuable and may advance our understanding of the interpersonal and motivational benefits of being an underdog (e.g., Goldschmied & Vandello, 2009; Lount, Pettit, & Doyle, 2016; Nurmohamed, 2014; Paharia, Keinan, Avery, & Schor, 2011).

An important likely boundary condition for our findings relates to how costly it is to reduce persistence. We acknowledge that reducing persistence, and quitting in particular, often come with economic and social costs. From an economic perspective, reducing persistence decreases (or eliminates) one’s likelihood of success. For example, if a tennis player quits, he forgoes his opportunity to win a match; if he persists, he still has some chance of coming back. Also, quitting is often socially stigmatized (Eriksson, Mao, & Villeval, 2015; Fershtman & Gneezy, 2011). Therefore, in some contexts, quitting may simply be too costly for individuals to pursue as an image management strategy. Notwithstanding, subtle forms of quitting, such as reallocating effort from one task to another, are common and unlikely to produce significant backlash, especially when work is difficult to observe and alternative tasks, even when they are less promising, are still valuable to pursue.

Further, we propose that prior quitting research has overlooked image management benefits that can, and often do, outweigh the costs of quitting. By quitting, people can withdraw from situations that make them feel embarrassed and ashamed, avoid being held accountable for poor outcomes, and have the opportunity to re-build their confidence and public image by excelling elsewhere. In contrast, completely failing and being identified as a "loser" can have significant psychological and social costs. The experience of failure can threaten self-image and cause individuals to perform less effectively on a subsequent task, especially for individuals with a strong prior reputation (Marr & Thau, 2014). In fact, merely anticipating failure or a loss of status can prompt individuals to take preemptive action (Chen et al., 2011; Pettit, Yong, & Spataro, 2010). In our investigation, we find that the prospect of losing, made salient by an initial setback, can prompt individuals to reduce persistence especially when losing violates expectations. These costs can sometimes be avoided if people reduce persistence when they encounter initial setbacks, assuming they anticipate a greater expectancy violation should they persist. 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). Tennis players in our field study could quit mid-match citing a physical injury; participants in our first laboratory experiment could justify their decision to switch topics by claiming that the initial trivia topics covered their weakest knowledge areas. In other contexts, people can point to a myriad of plausible reasons for reducing persistence, such as family commitments. We expect the image benefits of quitting to be moderated by the credibility of the justifications people provide. For example, executives who step down citing the need to spend time with family will be more credible when they have young children at home than when they do not. Potentially interesting avenues for future research include exploring how people balance the image management benefits and costs of quitting and how the perceived legitimacy of a quitting excuse moderates the relationship we document between performance expectations and persistence.

In all of our studies, individuals were aware that they were under public scrutiny, because their performance was either publicly announced as the task unfolded or became public at its conclusion. This aligns with the public observability of performance outcomes in many organizational contexts (e.g., periodic updates on product developments, posting sales rankings) and competitive environments (e.g., promotions, elections). We suspect that high expectations may reduce persistence even when people's performance is private because people care about maintaining a positive self-image. A potentially interesting line of future inquiry would be to compare the effects of performance expectations on persistence in public versus private settings and separate out the roles played by self-image concerns and impression management concerns.

Our investigation focused on the persistence decisions of individual decision makers. However, organizations may experience analogous situations where image management concerns play a role in persistence decisions. Zavyalova, Pfarrer, Reger, and Hubbard (2016) show that when a negative event befalls an organization with an outstanding reputation, this is a greater violation of expectations than it would be had that same setback affected a less prestigious organization. Further, organizations employ image management tactics when they anticipate negative expectancy violations (Graffin, Haleblian, & Kiley, 2016). Thus, high performance expectations may sometimes be a burden to organizations just as they can be to individuals, leading organizations to reduce their persistence in the face of adversity to avoid further reputational damage. Future work should extend our investigation to organizational persistence decisions (George et al., 2016). Though prior work has identified high performance expectations as an asset, high expectations can also be a liability when it comes to persistence.

REFERENCES

Ai, C., & Norton, E.C. 2003. Interaction terms in logit and probit models. ***Economics Letters***, 80: 123-129.

Angrist, J. D., & Pischke, J. S. 2009. ***Mostly harmless econometrics: An empiricist's companion.*** Princeton, NJ: Princeton University Press.

Apsler, R. 1975. Effects of embarrassment on behavior toward others. ***Journal of Personality and Social Psychology***, 32: 145-153.

Babad, E. Y., Inbar, J., & Rosenthal, R. 1982. Pygmalion, Galatea, and the Golem: Investigations of biased and unbiased teachers. ***Journal of Education Psychology,*** 74(4): 459-474.

Bandura, A. 1982. Self-efficacy mechanism in human agency. ***American Psychologist,*** 37(2): 122-147.

Bandura, A. 1986. ***Social foundations of thought and action: A social cognitive theory.*** Eaglewood Cliffs, NJ: Prentice-Hall.

Bandura, A. 1997. ***Self-efficacy: The exercise of control.*** New York: Freeman.

Barankay, I. 2012. ***Rank incentives: Evidence from a randomized workplace Experiment.*** Working paper, Wharton Business School, Philadelphia, PA.

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(6): 1447-1457.

Berger, J., & Pope, D. 2011. Can losing lead to winning? ***Management Science***, 57 (5): 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*(2): 200-206.

Berman, S. L., Down, J., & Hill, C. W. L. 2002. Tacit knowledge as a source of competitive advantage in the National Basketball Association. ***Academy of Management Journal***, 45 (1): 13-31.

Bernstein, E. S., & Li, S. X. 2016. ***Seeing where you stand: From performance feedback to performance transparency.*** Working paper, Harvard Business School, MA.

Breznik, K., & Batagelj, V. 2012. Retired matches among male professional tennis players. ***Journal of Sports Science and 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.

Bull, C., Schotter, A., & Weigelt, K. 1987. Tournaments and pieces rates: An experimental study. ***Journal of Political Economy,*** 95(1): 1-33.

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. 2003. Stereotype threat: the effect of expectancy on performance. ***European Journal of Social Psychology,*** 33(2): 267-285.

Chapman, G. B., & McCauley, C. 1993. Early career achievements of National Science Foundation (NSF) graduate applicants: Looking for Pygmalion and Galatea effects on NSF winners. ***Journal of Applied Psychology***, 78(5): 815-820.

Charness, G., Masclet, D., & Villeval, M. C. 2014. The dark side of competition for status. ***Management Science,*** 60(1): 38-55.

Chen, G., & Kilmoski, J. 2003. The impact of expectations on newcomer performance in teams as mediated by work characteristics, social exchanges, and empowerment. ***Academy of Management Journal,*** 46(5): 591-607.

Chen, M.K., & J.M. Shapiro. 2007. Do harsher prison conditions reduce recidivism? A discontinuity-based approach. ***American Law and Economics Review,*** 9(1): 1-29.

Chen, H., Ham, S. H., & Lim, N. 2011. Designing multiperson tournaments with asymmetric contestants: An experimental study. ***Management Science,*** 57(5): 864-883.

Conlon, E. J. 1980. Feedback about personal and organizational outcomes and its effect on persistence of planned behavioral changes. ***Academy of Management Journal,*** 23(2): 267-286.

Cotton, R. D., Shen, Y., & Livne-Tarandach, R. 2011. On becoming extraordinary: The content and structure of the developmental networks of Major League Baseball Hall of Famers. ***Academy of Management Journal***, 54 (1): 15-46.

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(3): 386-398.

de Hooge, I. E., Breugelmans, S. M., & Zeelenberg, M. 2008. Not so ugly after all: When shame acts as a commitment device. ***Journal of Personality and Social Psychology,*** 95(4): 933-943.

de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. 2010. Restore and protect motivations following shame. ***Cognition and Emotion,*** 24(1): 111-127.

Eden, D. 1990. ***Pygmalion in management: Productivity as self-fulfilling prophecy.*** Lexington, MA, England: Lexington Books.

Eden, D. 1992. Leadership and expectations: Pygmalion effects and other self-fulfilling prophecies in organizations. ***The Leadership Quarterly,*** 3(4): 271-305.

Eden, D., & Aviram, A. 1993. Self-efficacy training to speed reemployment: Helping people to help themselves. ***Journal of Applied Psychology***, 78(3): 352-360.

Eden, D., & Zuk, Y. 1995. Seasickness as a self-fulfilling prophecy: raising self-efficacy to boost performance at sea. ***Journal of Applied Psychology***, 80(5), 628.

Eriksson, T., Mao, L., & Villeval, M. C. 2015. ***Saving face and group identity.*** http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2613976.

Ewusi-Mensah, K., & Przasnyski, Z. H. 1991. On information systems project abandonment: An exploratory study of organizational practices. ***MIS Quarterly,*** 15(1): 67-86.

Fershtman, C., & Gneezy, U. 2011. The tradeoff between performance and quitting in high power tournaments. ***Journal of the European Economic Association***, 9(2): 318-336.

Finch, J. F., & Cialdini, R. B. 1989. Another indirect tactic of (Self-) image management: Boosting. ***Personality and Social Psychology Bulletin,*** 15(2): 222-232.

Flammer, C. 2015. Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. ***Management Science,*** 61(11): 2549-2568.

Frijda, N. H., Kuipers, P., & ter Schure, E. 1989. Relations among emotions, and emotional action readiness. ***Journal of Personality and Social Psychology,*** 57(2): 212-218.

Gelman, A., & Imbens, G. 2014. ***Why high-order polynomials should not be used in regression discontinuity designs.*** Working paper no. 20405, National Bureau of Economic Research, Cambridge, MA.

George, G., Dahlander, L., Graffin, S. D., & Sim, S. 2016. Reputation and status: Expanding the role of social evaluations in management research. ***Academy of Management Journal,*** 59(1): 1-13.

Ghose, A., Ipeirotis, P., & Li, B. 2013. Examining the impact of ranking on consumer behavior and search engine revenue. ***Management Science***, 60(7): 1632-1654.

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(4): 750-783.

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.

Gold, M.A. 1999. ***Pygmalion in cyberspace: Leaders' high expectancies for subordinate performance conveyed electronically versus face-to-face.*** Unpublished doctoral dissertation, The University at Albany, State University of New York.

Goldschmied, N., & Vandello, J. A. 2009. The advantage of disadvantage: Underdogs in the political arena. ***Basic and Applied Social Psychology***, 31(1): 24-31.

Graffin, S., Haleblian, J., & Kiley, J. T. 2016. Ready, AIM, acquire: Impression offsetting and acquisitions. ***Academy of Management Journal,*** 59(1): 232-252.

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(4): 419-434.

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(5): 1136-1145.

Hahn, J., Todd, P., & Van der Klaauw, W. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. ***Econometrica,*** 69(1): 201-209.

Imbens, G., & Kalyanaraman, K. 2009. ***Optimal bandwidth choice for the regression discontinuity estimator.*** Working paper no. 14726. National Bureau of Economic Research, Cambridge, MA.

Imbens, G., & Kalyanaraman, K. 2012. Optimal bandwidth choice for the regression discontinuity estimator. ***Review of Economic Studies,*** 79(3): 933-959.

Imbens, G., & Lemieux, T. 2008. Regression discontinuity designs: A guide to practice. ***Journal of Econometrics,*** 142(2): 615–635.

Isaac, M. S., & Schindler, R. M. 2014. The top-ten effect: Consumer's subjective categorization of ranked lists. ***Journal of Consumer Research,*** 40(6): 1181-1202.

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, UK: Oxford University Press.

Jones, E. E., Gergen, K. J., & Jones, R. G. 1963. Tactics of ingratiation among leaders and subordinates in a status hierarchy. ***Psychological Monographs: General and Applied***, 77(3): 1.

Keltner, D., & Buswell, B. N. 1997. Embarrassment: Its distinct form and appeasement functions. ***Psychological Bulletin,*** 122(3): 250-270.

Kenny, D. A., Kashy, D. A., & Cook, W. L. 2006. ***Dyadic data analysis.*** New York: Guilford Press.

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(5): 943-969.

Kolditz, T. A., & Arkin, R. M. 1982. An impression management interpretation of the self-handicapping strategy. ***Journal of Personality and Social Psychology***, 43(3): 492-502.

Kray, L. J., Thompson, L., & Galinsky, A. 2001. Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. ***Journal of Personality and Social Psychology,*** 80(6): 942-958.

Larrick, R. P., Timmerman, T. A., Carton, A. M., & Abrevaya, J. 2011. Temper, temperature, and temptation: Heat-related retaliation in baseball. ***Psychological Science,*** 22(4): 423-428.

Leary, M. R., Barnes, B. D., & Griebel, C. 1986. Cognitive, affective, and attributional effects of potential threats to self-esteem. ***Journal of Social and Clinical Psychology,*** 4: 461-474.

Leary, M. R., & Kowalski, R. M. 1990. Impression management: A literature review and two-component model. ***Psychological bulletin***, 107(1): 34.

Lee, D. S., & Lemieux, T. 2010. Regression discontinuity designs in economics. ***Journal of Economic Literature,*** 48(2): 281-355.

Legge, S., & Schmid, L. 2014. ***Rankings, success, and individual performance: Evidence from a natural experiment.*** Working Paper no. 2013-13, U. of St. Gallen Law & Economics, St. Gallen, Switzerland.

Lim, N. 2010. Social loss aversion and optimal contest design. ***Journal of Marketing Research,*** 47(4): 777-787.

Lim, N., Ahearne, M. J., & Ham, S. H. 2009. Designing sales contests: Does the price structure matter? ***Journal of Marketing Research,*** 46(3): 356-371.

Lindsay, D. 2015. ***Livestrong taps former Komen executive as new CEO,*** media release, 24 March, The Chronicle of Philanthropy, < https://philanthropy.com/article/Livestrong-Taps-Former-Komen/228749>. viewed 16 May 2016.

Lount, R. B., Jr., Pettit, N. C., & Doyle, S. P. 2016. ***Motivating underdogs and favorites.*** Working paper, Ohio State University, Columbus, OH.

Luca, M. 2014. ***Reviews, reputation, and revenue: The case of Yelp.com.*** Working paper no. 12-016, Harvard Business School, Cambridge, MA.

Luca, M., & Smith, J. 2013. Salience in quality disclosure: Evidence from the US News college rankings. ***Journal of Economics & Management Strategy***, 22(1): 58-77.

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(1): 223-248.

McCambridge, R. 2016. ***An ominous and sudden departure at the Livestrong Foundation,*** media release, 21 January, Nonprofit Quarterly, <https://nonprofitquarterly.org/2016/01/21/an-ominous-and-sudden-departure-at-the-livestrong-foundation/>. viewed 16 May 2016.

McNatt, D. B. 2000. Ancient Pygmalion joins contemporary management: A meta-analysis of the result. ***Journal of Applied Psychology,*** 85(2): 314-322.

McNatt, D. B., & Judge, T. A. 2004. Boundary conditions of the Galatea effect: A field experiment and constructive replication. ***Academy of Management Journal***, *47*(4): 550-565.

Merton, R. K. 1948. The self-fulfilling Prophecy. ***The Antioch Review,*** 8(2): 193-210.

Meyers, J. 2015. ***The 10 real reasons Harry Reid quit the senate,*** media release, 27 March, Newsmax, <http://www.newsmax.com/TheWire/harry-reid-quit-senate-majority-leader/2015/03/27/id/634915/>. viewed 16 May 2016.

Miller, R. S. 1992. The nature and severity of self-reported embarrassing circumstances. ***Personality and Social Psychology Bulletin,*** 18(2): 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: Russell Sage Foundation.

Miller, R. S., & Tangney, J. P. 1994. Differentiating embarrassment and shame. ***Journal of Social and Clinical Psychology,*** 13(3): 273-287.

Mills, M. P., & Mills, M. A. 2014. ***Every datum tells a story,*** media release, 29 June, City Journal, < http://www.city-journal.org/html/every-datum-tells-story-11392.html>. viewed 16 May 2016.

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,*** 53(4): 701-722.

Modigliani, A. 1971. Embarrassment, facework, and eye contact: Testing a theory of embarrassment. ***Journal of Personality and Social Psychology,*** 17(1): 15-24.

Mosher, D. L., & White, B. B. 1981. On differentiating shame and shyness. ***Motivation and Emotion,*** 5(1): 61-74.

Nurmohamed, S. 2014. Over or under? The motivational implications of an underdog image. Unpublished doctoral dissertation, University of Michigan.

O'Connor, K. M., & Arnold, J. A. 2001. Distributive spirals: Negotiation impasses and the moderating role of disputant self-efficacy. ***Organizational Behavior and Human Decision Processes,*** 84(1): 148-176.

Oz, S., & Eden, D. 1994. Restraining the Golem: Boosting performance by changing the interpretation of low scores. ***Journal of Applied Psychology,*** 79(5): 744-754.

Paharia, N., Keinan, A., Avery, J., & Schor. J. B. 2011. The underdog effect: The marketing of disadvantage and determination through brand biography. ***Journal of Consumer Research,*** 37(5): 775-790.

Petkova, A. P., Wadhwa, A., Yao, X., & Jain, S. 2014. Reputation and decision making under ambiguity: A study of US venture capital firms' investments in the emerging clean energy sector. ***Academy of Management Journal,*** 57(2): 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(8): 1579-1584.

Pettit, N. C., Yong, K., & Spataro, S. E. 2010. Holding your place: Reactions to the prospects of status gains and losses. ***Journal of Experimental Social Psychology,*** 46(2): 396-401.

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(3): 359-374.

Pope, D. 2009. Reacting to rankings: Evidence from "America's Best Hospitals". ***Journal of Health Economics,*** 28(6): 1154-1165.

Pope, D., & Schweitzer, M. E. 2011. Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. ***American Economic Review,*** 101: 129-157.

Reifman, A. S., Larrick, R. P., & Fein, S. 1991. Temper and temperature on the diamond: The heat-aggression relationship in major-league baseball. ***Personality and Social Psychology Bulletin***, 17: 580-585.

Rhodewalt, F., Saltzman, A. T., & Wittmer, J. 1984. Self-handicapping among competitive athletes: The role of practice in self-esteem protection. ***Basic and Applied Social Psychology,*** *5*(3): 197-209.

Rosenthal, R., & Jacobson, L. 1968a. Teacher expectations for the disadvantaged. ***Scientific American,*** 218(4): 19-23.

Rosenthal, R., & Jacobson, L. 1968b. ***Pygmalion in the classroom: Teacher expectation and pupils' intellectual development.*** New York: Holt, Rinehart & Winston

Schrank, W. R. 1968. The labeling effect of ability grouping. ***The Journal of Educational Research,*** 62(2): 51-52.

Song, H., Tucker, A. L., Murrell, K. L., & Vinson, D. R. 2015. ***Public relative performance feedback in complex service systems: Improving productivity through the adoption of best practices.*** Harvard Business School Research Paper Series, no. 16-043.

Stevens, C. K., & Kristof, A. L. 1995. Making the right impression: A field study of applicant impression management during job interviews. ***Journal of Applied Psychology***, 80(5): 587-606.

Steele, C. M. 1988. The psychology of self-affirmation: Sustaining the integrity of the self. ***Advances in Experimental Social Psychology,*** 21: 261-302.

Steele, C. M. 1997. A threat in the air: How stereotypes shape intellectual identity and performance. ***American Psychologist,*** 52(6): 613-629.

Steele, C. M., & Aronson, J. 1995. Stereotype threat and the intellectual test performance of African Americans. ***Journal of Personality and Social Psychology,*** 69(5): 797-811.

Swann, W. B., & Hill, C. A. 1982. When our identities are mistaken: Reaffirming self-conceptions through social interaction. ***Journal of Personality and Social Psychology,*** 43(1): 59-66.

Tangney, J. P. 1999. The self-conscious emotions: Shame, guilt, embarrassment and pride. In T. Dalgleish & M. J. Power (Eds.), ***Handbook of cognition and emotion***: 541-568. New York: John Wiley & Sons Ltd.

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(6): 1256-1269.

Tennis Industry Association. 2013. ***State of the industry 2013 edition***. <https://www.tennisindustry.org/pdfs/TIA_RSI_News_June%20.pdf>. viewed 23 May 2016.

Weber, M., & Camerer, C. F. 1998. The disposition effect in securities trading: an experimental analysis. ***Journal of Economic Behavior and Organization,*** 33(2): 167-184.

Weinberg, R. S., Yukelson, D., & Jackson, A. 1980. Effect of public and private efficacy expectations on competitive performance. ***Journal of Sport Psychology***, 2(4): 340-349.

Whyte, G., & Saks, A. M. 2007. The effects of self-efficacy on behavior in escalation situations. ***Human Performance,*** 20(1): 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(5): 415-432.

Wooldridge, J. M. 2010. ***Econometric Analysis of Cross Section and Panel Data.*** Boston: The MIT Press.

Zavyalova, A., Pfarrer, M. D., Reger, R. K., & Hubbard, T. D. 2016. Reputation as a benefit and a burden? How stakeholder's organizational identification affects the role of reputation following a negative event. ***Academy of Management Journal,*** 59(1): 253-276.

TABLES AND FIGURES

Table 1. Descriptive statistics

Variable	Mean	Std. Dev.
Quit	0.014	0.117
Favorite	0.50	0.50
Lost First Set	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 ($)	245374.20	660006.90
Year	1998.83	9.92
Round	1.92	1.13
Win the Match	0.50	0.50

Note. Descriptive statistics are calculated for 657,030 player-match-observations from 328,515 matches with the exception of *lost first set* 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 Appendix B.

Table 2. Ordinary least squares regression-discontinuity models predicting quitting mid-match (Models 1-3), winning (Models 4-6), and quitting prior to the start of a match (Model 7) as a function of whether a player was the favorite to win a match while controlling for the player's skill, his opponent's skill, his skills relative to his opponent's, and other observable player and match characteristics.

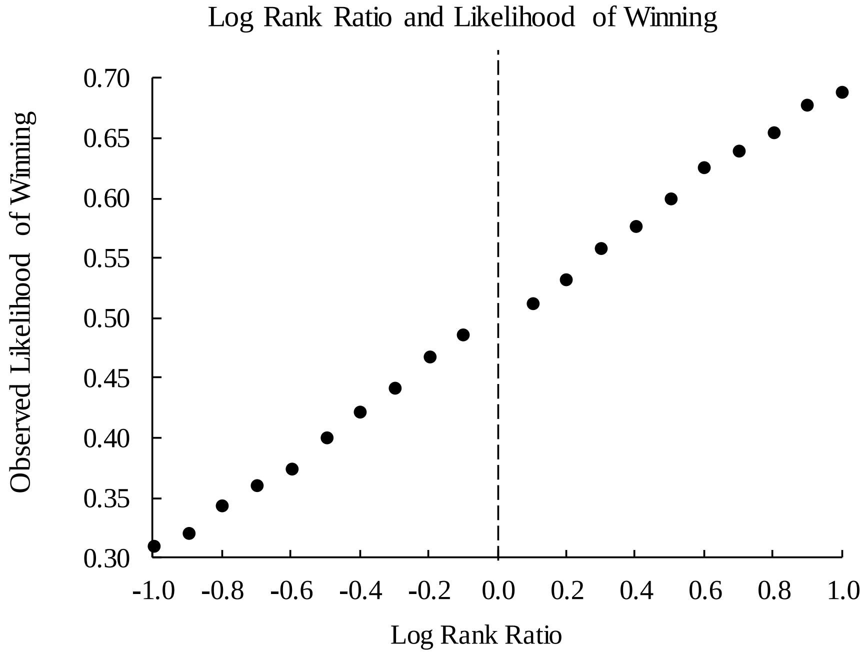
Outcome Variable	Quitting Mid-Match			Winning the Match			Quitting Prior to a Match
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Favorite Indicator	0.0037***	0.0002	-0.0005	-0.0027	-0.0028	-0.0021	-0.0001
	(0.0010)	(0.0005)	(0.0005)	(0.0027)	(0.0028)	(0.0028)	(0.0002)
Lost First Set x Favorite Indicator			0.0054***			0.0001	
			(0.0009)			(4.58E-05)	
Lost First Set Indicator			0.0294***			-0.6441***	
			(0.0019)			(0.0034)	
Rank	0.0061	-0.0020	0.0021	0.0298*	0.0297*	0.0132	-0.0003
	(0.0043)	(0.0022)	(0.0022)	(0.0130)	(0.0120)	(0.0105)	(0.0009)
Opponent's Rank	-0.0061	0.0020	-0.0021	-0.0298*	-0.0297*	-0.0132	0.0003
	(0.0043)	(0.0022)	(0.0022)	(0.0130)	(0.0120)	(0.0105)	(0.0009)
Log(Rank Ratio)	0.0006	-0.0003	-0.0010**	0.0820***	0.0820***	0.0793***	0.0002
	(0.0006)	(0.0004)	(0.0004)	(0.0020)	(0.0021)	(0.0018)	(0.0001)
Log(Rank Ratio)^2	-0.0003	0.0001	0.0007***	0.0101***	-0.0102***	-0.0091***	0.0002***
	(0.0002)	(0.0001)	(0.0001)	(0.0008)	(0.0008)	(0.0005)	(0.0001)
Log(Rank Ratio)^3	4.48E-06	-0.0001	2.38E-05	-0.0005*	-0.0005**	-0.0008***	-1.41E-05
	(4.32E-05)	(4.82E-05)	(3.40E-05)	(0.0002)	(0.0002)	(0.0001)	(9.96E-06)
Log(Rank Ratio)^4	-2.37E-06	7.88E-06	-2.47E-05**	-0.0002***	0.0002***	0.0002***	-8.94E-06***
	(1.14E-05)	(1.14E-05)	(8.53E-06)	(4.56E-05)	(4.47E-05)	(2.86E-05)	(2.66E-06)
Lost First Set X Log(Rank Ratio)			0.0024***			4.07E-05	
			(0.0006)			(3.01E-05)	
Lost First Set X Log(Rank Ratio)^2			-0.0011***			0.0181***	
			(0.0002)			(0.0011)	
Lost First Set X Log(Rank Ratio)^3			-8.82E-05*			-3.05E-06	
			(4.25E-05)			(2.25E-06)	
Lost First Set X Log(Rank Ratio)^4			2.73E-05**			-0.0004***	
			(1.00E-05)			(5.74E-05)	
Player's Age	0.0022***	0.0005	0.0014***	-0.0045**	-0.0021	-0.0032**	0.0003*
	(0.0005)	(0.0003)	(0.0003)	(0.0015)	(0.0018)	(0.0011)	(0.0001)
Player's Age^2	-3.35E-05***	-6.90E-06	-2.00E-05***	5.24E-05^	1.40E-05	3.31E-05	-3.67E-06
	(9.15E-06)	(6.53E-06)	(5.58E-06)	(2.79E-05)	(3.45E-05)	(2.12E-05)	(2.39E-06)
Prize Money	1.36E-09*	7.22E-10^	1.04E-09*	1.06E-09	-8.27E-10	1.03E-10***	-3.76E-11
	(6.52E-10)	(4.14E-10)	(4.07E-10)	(2.06E-09)	(1.94E-09)	(2.31E-11)	(1.38E-10)
Player's Rank, Fully Interacted with Log(Rank Ratio) Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Opponent's Rank, Fully Interacted with Log(Rank Ratio) Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Match Surface Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tour Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tournament Round Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tour x Round Fixed Effect Interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Set Score Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	No
Matches Included in Analysis	Those in which the target player lost the first set	Those in which the target player won the first set	Those in which the first set was completed	Those in which the target player lost the first set	Those in which the target player won the first set	Those in which the first set was completed	All
Observations	323,683	323,683	647,366	323,683	323,683	647,366	655,774
R-squared	0.0057	0.0017	0.0072	0.0555	0.0553	0.4466	0.0016

Note. This table reports analyses that exclude matches where two players had identical rankings. Bootstrapped standard errors are reported in parentheses. P-values are calculated based on bootstrapped standard errors.

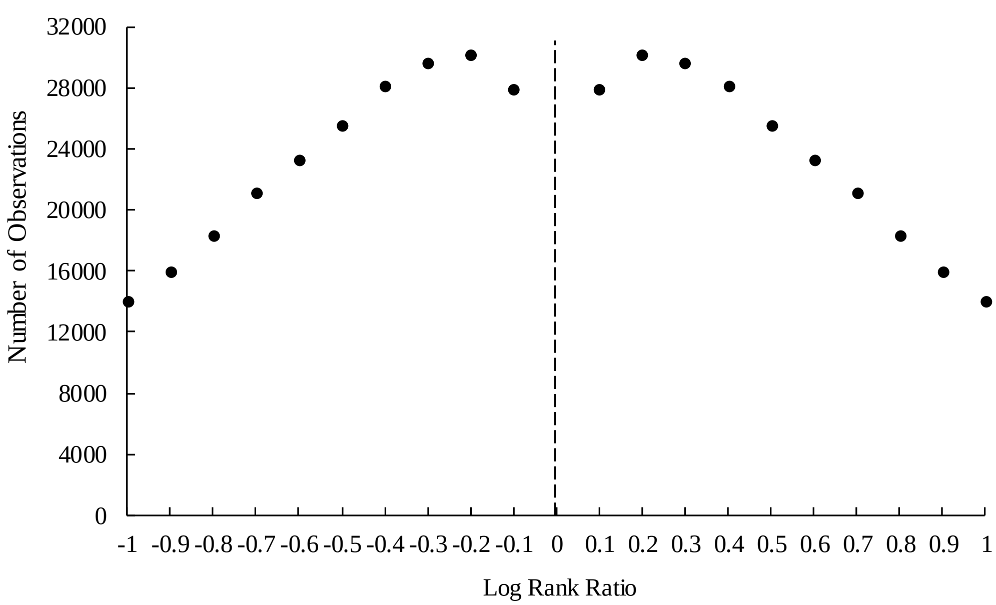
^, *, **, *** denotes significance at the 10%, 5%, 1%, and 0.1% level, respectively

Figure 1. Plots of the likelihood of winning a match as a function of the target player's *log rank ratio* (Panel A) and the distribution of *log rank ratio* (Panel B)

Panel A. The relationship between a target player's *log rank ratio* relative to his opponent and the target player's likelihood of winning the match.

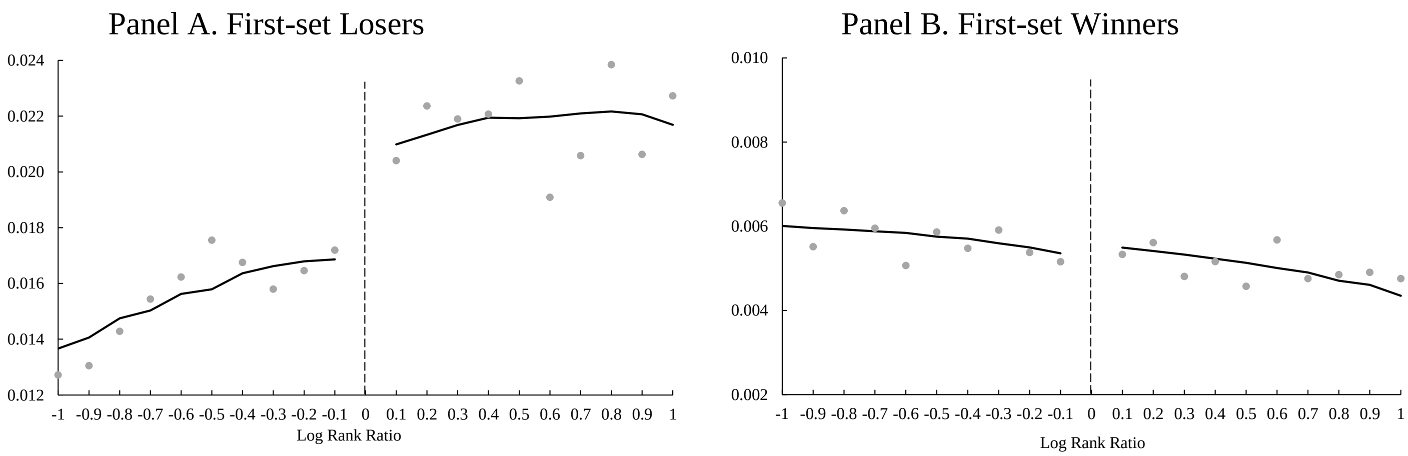


Panel B. Distribution of target players' *log rank ratio*



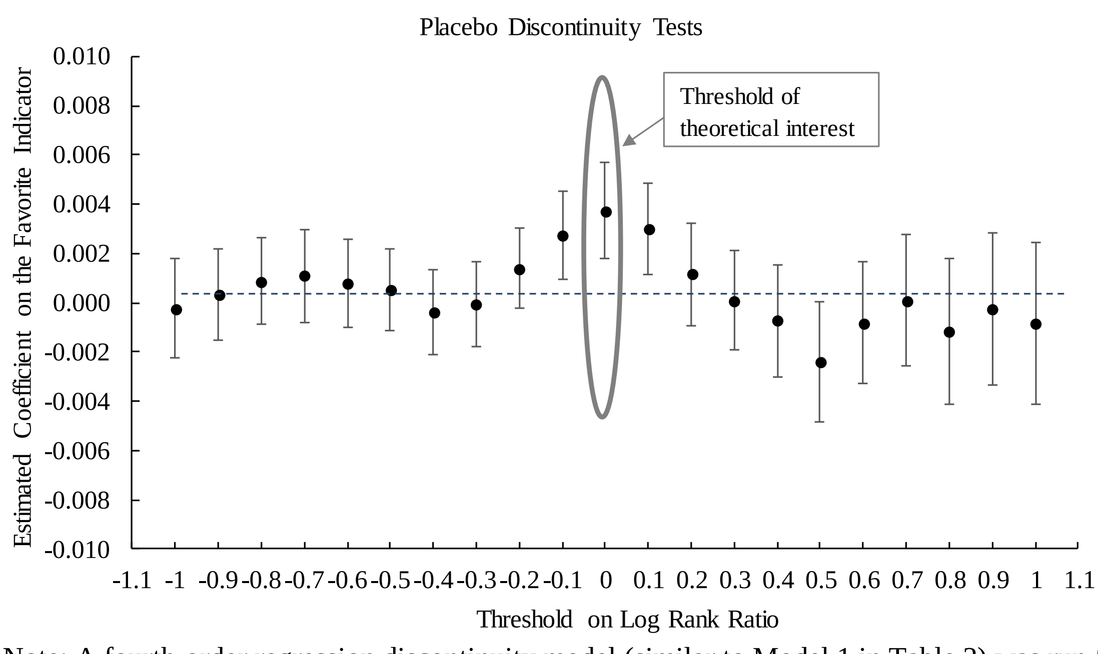
Note. We depict the 50% of matches that fall closest to the underdog-favorite threshold. For each of the 20 bins of width 0.1 of the *log rank ratio* variable, the observed likelihood of winning is calculated as the average likelihood of winning across all observations in that bin.

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.



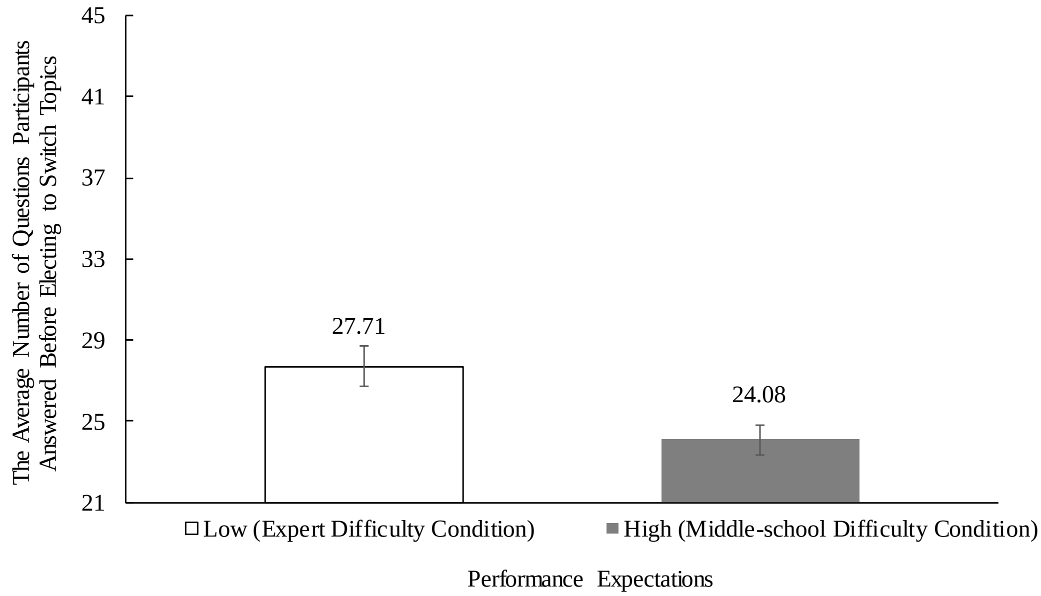
Note. We depict the 50% of matches that fall closest to the favorite-underdog threshold. Observations are grouped into bins of width 0.1 based on the *log rank ratio*. Panel A shows outcomes for first set losers; Panel B for first set winners. Each dot represents the probability that a player quit mid-match when that player's *log rank ratio* fell in a particular bin. 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.

Figure 3. Estimated change in quitting at different (placebo) regression discontinuity thresholds as well as the threshold of theoretical interest.



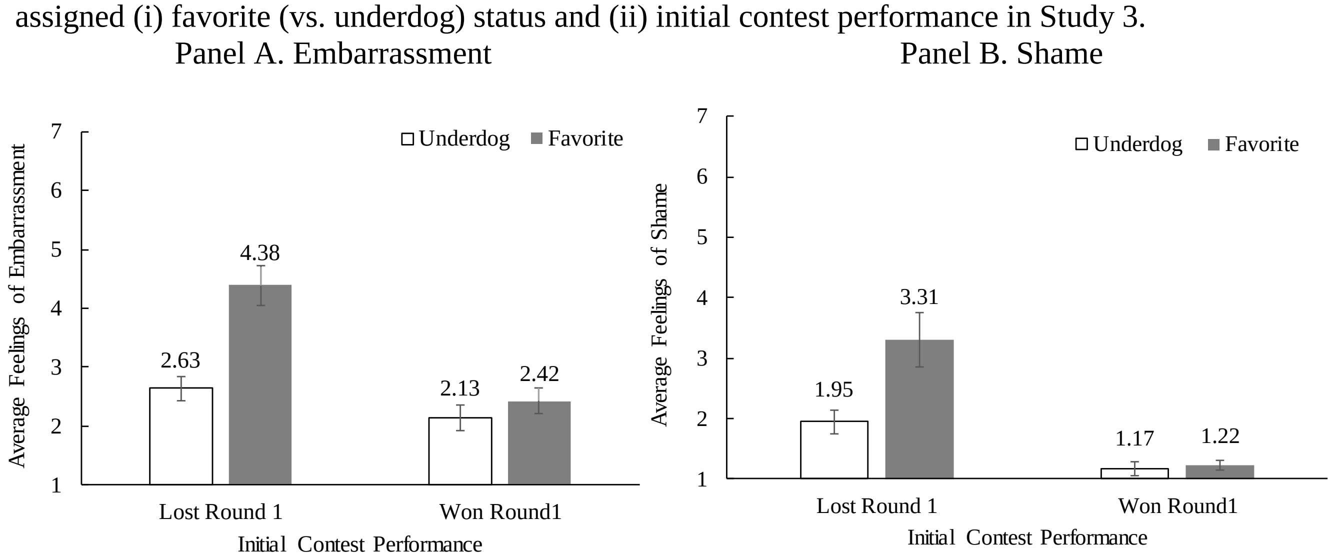
Note: A fourth-order regression discontinuity model (similar to Model 1 in Table 2) was run for each threshold value of log rank ratio between −1.0 and 1.0 at 0.1 intervals. Each dot represents the coefficient on the favorite indicator for a given threshold. Error bars represent 95% confidence intervals.

Figure 4. Persistence as a function of performance expectations in Study 2.



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

Figure 5. Feelings of embarrassment (Panel A) and shame (Panel B) as a function of randomly assigned (i) favorite (vs. underdog) status and (ii) initial contest performance in Study 3.



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