UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

FACULTY OF BUSINESS ADMINISTRATION AND ECONOMICS

Working Paper Series

Working Paper No. 2022-15

Are women more resilient? Gender differences in the reaction to negative feedback

Lisa Beck-Werz and Thomas Fritz

July 2022

wiwi

Are women more resilient?
Gender differences in the reaction to negative feedback

Lisa Beck-Werz[a,b,*], Thomas Fritz[b]

[a] *University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany*
[b] *Aachen University of Applied Sciences, Bayernallee 11, 52066 Aachen, Germany*

[*] Corresponding author
E-mail address: beck-werz@fh-aachen.de
Phone: +49 175 3189430

Declaration of interests: none

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Although the gender gap in labor markets is steadily narrowing, there is a persistent shortage of women in competitive high-ranking positions. A large body of literature suggests that gender differences in preference for and reaction to competition are potential explanations for the scarcity of women available for leadership positions. Recent studies indicate through laboratory experiments that positive feedback eliminates gender differences in competitiveness and self-confidence, whereas negative feedback and bad experiences cause women to be more likely to withdraw from competition and reduce performance. Our study reveals a new facet of this phenomenon: Women who voluntarily participate in competitive settings are more likely to return to competition after a negative experience than men. While this underscores that feedback indeed causes different reactions in men and women, it suggests that the characteristics of the feedback is decisive: Men tend to avoid competitive situations after receiving implicit negative feedback and increasingly compete after positive experiences. Women, by contrast, re-enter a competition regardless of the nature of the previous experience.

Keywords: competition; gender differences; negative feedback; reaction to failure; career decisions; gender gap

JEL Codes: J16, D91, M50

1. Introduction

There is a substantial and persistent gender gap in labor market outcomes such as wage and share of women in senior positions (see e.g. Boll & Leppin, 2015; Goldin, 2014). For example, in 2019 the percentage of female CEOs of Fortune 500 firms peaked at 6,6% (Zillman, 2019). In Europe, women represent 33% of the STOXX Europe Companies' board members, but only 28 (4.7%) of the 600 CEO positions are held by women (EWOB, 2019).

A wide range of scientific contributions show that the gap can only partly be explained by sorting effects, discrimination, or human capital factors such as education and experience (Black & Strahan, 2001; Blau & Kahn, 2017; Card et al., 2016). A growing body of literature in experimental economics explains the gender gap in market labor outcomes with gender differences in preference for (Niederle & Vesterlund, 2007[1]) as well as reaction to competition (Gneezy et al., 2003[2]). These studies confirm the assumption that gender differences in competitivity carry over to different decisions in later education, career choices and finally career outcomes (Almås et al., 2016; Buser et al., 2014). Consequently, a significant part of recent contributions concentrate on the investigation of potential policy interventions and measures (Balafoutas & Sutter, 2012; Niederle et al., 2013). The overall objective is to enable women and men to „make choices that reflect their underlying preferences over outcomes rather than reflecting differences in psychological attributes which play a role due to the environment in which decisions are made“ (Niederle, 2016, p. 9).

A considerable body of previous literature in various fields found that feedback on relative performance can eliminate existing gender gaps in competitiveness (e.g. Wozniak et al., 2011), self-confidence (e.g. Lenney, 1977), ability to compete under pressure and performance (e.g. Cotton et al., 2013). Until now, only few contributions investigated gender-specific effects of negative feedback and setbacks. Two pioneer studies demonstrated in repeated laboratory experiments that after a loss women lower both effort (Gill & Prowse, 2014) and performance, whereas men react to loss by seeking a more challenging target (Buser, 2016).

[1] See also Balafoutas and Sutter (2019); Dohmen and Falk (2011); Flory et al. (2018); Kamas and Preston (2012); Sutter and Glätzle-Rützler (2010)

[2] See also Antonovics et al. (2009); Gill and Prowse (2014); Gneezy and Rustichini (2004)

Educational and career paths are very similar to repeated competitions: in both, participants repeatedly experience and respond to successes and losses; in both, repeatedly decisions on whether and how to (re)enter a competitive situation are necessary. Students decide to study, to obtain degrees or to enter the labor market - employees decide whether to engage in competition for job offers, promotions and salary increases (Bognanno, 2001; Eriksson, 1999). Different choices made at an early stage of the career help to explain some of the gender gap on the supply side of the labor market.

The aim of this study is to better understand how women and men respond to setbacks and negative feedback in a natural environment. The results are intended to serve as a basis for further investigations of potential policies and institutional measures that can be used to dissolve the gender gap in labor market outcomes. In our project we analyzed the behavior of marathon runners who choose a too fast race pace and receive implicit negative feedback which is manifested by a severe slowdown during the race ("hit the wall").

Our research is structured along two questions:

A) **Impact of negative feedback on competitiveness / preference to participate in further competition**
B) **Impact of negative feedback on behavior in subsequent competitive situations**

A distinct advantage of the sport context and the competitive environment of marathon races is the self-selection of participants. Various studies show considerably different preferences and behaviour among a professional and managerial population (see e.g. Croson & Gneezy, 2009 for an overview). For example, among professional fund managers no gender difference regarding preference for risk could be detected (Atkinson et al., 2003). Therefore, the setting is especially suitable for drawing parallels to the labor market later on when discussing potential implications. In the labor market as well as in marathons (and in contrast to most laboratory settings), participants show a certain preference to compete and voluntarily join the competitive situation. Moreover, endurance sports, characterized by transparent and quantifiable results, allow for a particularly objective examination of behavior and its consequences.

1.1.Feedback and Gender Gap

The wide range of effects of feedback has been studied extensively in many previous studies. People increase effort levels and productivity when they are told how they perform relative to others (Clark & Oswald, 1996; Easterlin, 1995; Luttmer, 2005; Szymanski & Harkins, 1987; White et al., 1995). Kuhnen and Tymula (2012) demonstrated that people even put in more effort if they are promised relative feedback after a task, even if they are not additionally rewarded for the extra effort. Azmat and Iriberri (2010) showed in an university context that students who received information regarding their relative performance improved their grades across the distribution. Feedback helps people to better assess their performance in the future and leads to reduced positivity bias in judgment (see Kluger & DeNisi, 1996 for an overview). Additionally, feedback increases the overall level of self-confidence (Ehrlinger & Dunning, 2003; McCarty, 1986). Feedback plays a particular role in research of gender differences. A broad selection of studies exhibited that feedback can help to eliminate existing gender gaps in willingness to compete (e.g. Wozniak et al., 2011), self-confidence (e.g. Lenney, 1977), ability to compete under pressure and performance (e.g. Cotton et al., 2013). The following section provides an overview of previous research on reducing gender differences through feedback.

1.1.1. Feedback and the Gender Gap in Competitivity

A considerable body of previous research in various fields describes that women have a significant lower propensity to participate in competitive activities (see e.g. Croson & Gneezy, 2009; Gneezy & Rustichini, 2004; Niederle & Vesterlund, 2007). The gender gap in competitivity is particularly pronounced when tasks are not stereotypically female but described as challenging. These results persist even when controlling for risk aversion (see Niederle, 2016 for a detailed overview).

In 2011, Ertac and Szentes showed that these clear gender differences in competitive choices disappeared as soon as the participants received feedback on how they performed relative to others in the past (see also Balafoutas & Sutter, 2019; Masclet et al., 2015). Wozniak et al. (2011) confirmed the results and further illustrated in detail that relative performance feedback led high-ability women opting for competitive schemes, whereas low-ability men were more likely to pull out of the competition.

1.1.2. Feedback and the Gender Gap in Self-Confidence

Previous research has also extensively investigated and confirmed the lower self-confidence of women (see, e.g., Beyer & Bowden, 1997; Lundeberg et al., 1994). Different explanatory patterns can be found for this phenomenon. Women are more susceptible to negative self-assessments and attribute success to external factors (e.g., luck) and failure to internal factors (e.g., lack of talent), whereas men do the opposite (Dweck et al., 1978). Lundeberg et al. suggested: "The typical perception of women's lack of confidence, rather than men's overconfidence, may be the result of comparing prospective general confidence rather than retrospective and task or item-specific confidence" (1994, p. 120). Lenney (1977, p.11) noted in her comprehensive literature review that gender differences in the assessment of one's own performance can be reduced through clear feedback: when feedback is „unequivocal and immediately available, women do not make lower ability estimates than men. However, when such feedback is absent or ambiguous, women seem to have lower opinions of their abilities and often do underestimate relative to men." Kimball and Gray (1982) and Bench et al. (2015) confirmed the results and showed that detailed feedback about past performance helped to eliminate gender differences in the assessment of own performance across various academic tasks.

1.1.3. Feedback and the Gender Gap in Performance and Ability to Compete

Several contributions showed that men tend to outperform women in competitive situations (Gneezy et al., 2003; Gneezy & Rustichini, 2004). However, contrary results were also found in some studies which found no significant gender differences in competitive performance (Günther et al., 2010; Niederle & Vesterlund, 2007). Reasons for the different results could be that men appear significantly more motivated as soon as competitions involve stereotypical male-tasks (Günther et al., 2010). Moreover, numerous studies have shown that men are significantly more

ambitious and perform better when they are in direct competition with women (Antonovics et al., 2009; Gneezy et al., 2003; Gneezy & Rustichini, 2004; Price, 2008). Cotton et al. (2013) found that in repeated competitions where participants gained experience in the first round, no gender difference remained. Especially in terms of performance, women scored equally well in neutral tasks and outperformed men in later rounds of the tournament. Even after a two-week break, the male performance advantage did not reappear. The study of Cotton et al. (2013) also showed that dealing with pressure in competition was a reason for gender differences in the first round of the competition: when time pressure was removed, no significant gender gap was observable - even in the first round.

Generally, several studies indicated that women respond to other types of feedback than men. Women react strongly to feedback regarding how they perform relatively to others, whereas men are motivated by the competitive environment and general information about the levels of their opponents (Berlin & Dargnies, 2016). If women were informed about their absolute performance, without information about the performance of their opponents, gender differences in competitiveness could not be reduced (Dohmen & Falk, 2011).

1.2. Negative Feedback and the Reaction to Loss and Failure

Previous studies already explored the reaction to negative feedback and, in particular how individuals cope with success and failure. Most of the studies agree that (perceived) success improves and (perceived) failure reduces future performance (Bélanger et al., 2013; Gill & Prowse, 2012, 2014). Loss and failure in competitive situations had negative effects on future effort (Elliot & Church, 1997; Gill & Prowse, 2014), self-confidence, intrinsic motivation (McCarty, 1986) as well as performance (Legge & Schmid, 2013; Rosenqvist & Nordström Skans, 2015; Wozniak, 2011). In the context of sport, Haenni (2019) showed that losing led athletes to take a 10% longer break before the next competition, although the absence from competition had a negative effect on the individual ranking of the athlete.

First studies investigated gender differences in the reaction to loss and failure. Elliot and Thrash (2004) indicated in their psychological contribution that women are generally more influenced by fear of failure. Storek and Furnham (2014, p.48) suggested “it is possible that male hubris is less vulnerable to (negative) feedback […].” Gill and Prowse (2014) showed in a repeated, real

effort experiment that all participants basically reduced their effort after a loss, but men reduced their effort only when they failed to win a major prize. Buser (2016) studied the reaction to winning and losing in a two-round tournament laboratory experiment: participants completed an arithmetic task in the first round and were then informed about the result. In a second round the participants chose an individual performance target. The higher the target, the higher the possible win, but no money was given if the target was not reached. Buser (2016) was able to identify substantial gender differences: Men reacted to loss by seeking a more challenging target, whereby women responded to loss by reducing their performance.

Field and laboratory experiments in gender science show differing results and characteristics (cf. e.g. Antonovics et al., 2009). The aim of our study was to investigate the effect of negative feedback on men and women under real-life conditions by studying the behaviour of amateur marathon runners. The following chapter outlines the design of the study and in particular the sports context and environment of our investigations.

2. Study Design

Endurance contests such as marathons are one of the most genuine and severe judge you can face. Every athlete has to make two crucial decisision that are decisive for pain and pleasure in a race: 1) the fundamental decision to start the race (which includes the assessment that one can complete it, and 2) the choice of a suitable race pace, which is expressed in particular in the choice of the intial pace.

This unavoidable decision about the initial race pace is just as unavoidably provided with physical feedback as the race progresses. Athletes who choose an initial pace that is too fast inevitably have to slowdown in the second half of the race and receive immediate negative (and often painful) feedback on the assessment of their performance prior to the race. This negative feedback constitutes a drastic and impressive experience for the runner for several reasons. Firstly, due to the associated physical exhaustion and pain and secondly, the awareness of an inaccurate self-assessment causes a reduction of overall pleasure (McGraw et al., 2004). For this reason, runners' slowdown is an excellent metric to measure the implicit feedback runners experience during their race.

There is no tactically logical reason for an intentional variation of the race pace. There is a clear consensus in the existing literature that even or negative pacing[3] is the best racing strategy for endurance competitions that last several hours (see e.g. Abbiss & Laursen, 2008; March et al., 2011). Studies exploring different pacing strategies show that athletes with consistent pacing can deliver the best performance (Deaner et al., 2015). Regardless of what an athlete's ambition is - whether they want to run as fast as possible or whether they just want to finish the race - there is no motivation for an amateur athlete to start too fast or behave tactically in the first half of a marathon. The operationalization of a severe slowdown as a measure of negative feedback also has great advantages due to its objectivity and transparency. In contrast to participants in laboratory settings and artificial exercises, athletes receive feedback on a task that they have been dealing with intensively for months (and years) and have specifically trained for.

Marathon competitions are particularly suitable for the analysis. Due to the length of the course, the speed must be chosen even more carefully than in half marathons, as the effect of an overestimated initial speed is much more evident and painful. In contrast to ultra-marathons (distances over 50km), marathons are mass events and the large, heterogeneous data sets allow a comprehensive and differentiated analysis of the athletes.

The aim of our study is to explore the influence of negative feedback on marathon runners' behavior and, in particular, to investigate gender differences. Therefore, we compare the behavior of two groups of marathon runners:

Group „negative Feedback": Runners with a slowdown of more than 30% in the second half of the race[4]. A deceleration in pace on the second half marathon of >30%. has already been used in other studies as a threshold for significant slowdown (Deaner et al., 2015). In our initial sample, 4.1% of 77,878 marathon runners experienced such a severe slowdown. We can assume that athletes do not intentionally choose such a sharp deviation on the second half of the marathon and that

[3] Negative Pacing (or negative splitting) describes a pacing strategy that involves running the first half of a race slower than the second half.

[4] In the following, we define slowdown as $\frac{\text{pace 2n half marathon}}{\text{pace 1st half marathon}}$.

the undesired physical feedback causing the slowdown constitutes a considerable negative experience.

Group „no negative Feedback": Runners with a constant pace who deviate from the initial race pace by a maximum of 5% in the second half of the race. This applies to 39.0% of the 77,878 runners in the races considered.

Our analysis is structured along two main questions:

A. Impact of negative feedback on competitiveness / preference to participate in further competition

Re-entering a marathon race in the following year (year 2[5])

i) Does negative feedback affect the probability of re-entering a marathon in the following year?

ii) Are there gender-specific differences in the decision to re-enter?

B. Impact of negative feedback on behavior in subsequent competitive situations

Behavior in the race in year 2

i) Does negative feedback affect the racing behavior in year 2 and do runners adjust their competitive approach and racing strategy?

ii) Are there gender-specific differences in the adjustment of behavior in year 2?

3. Data and Descriptive Evidence

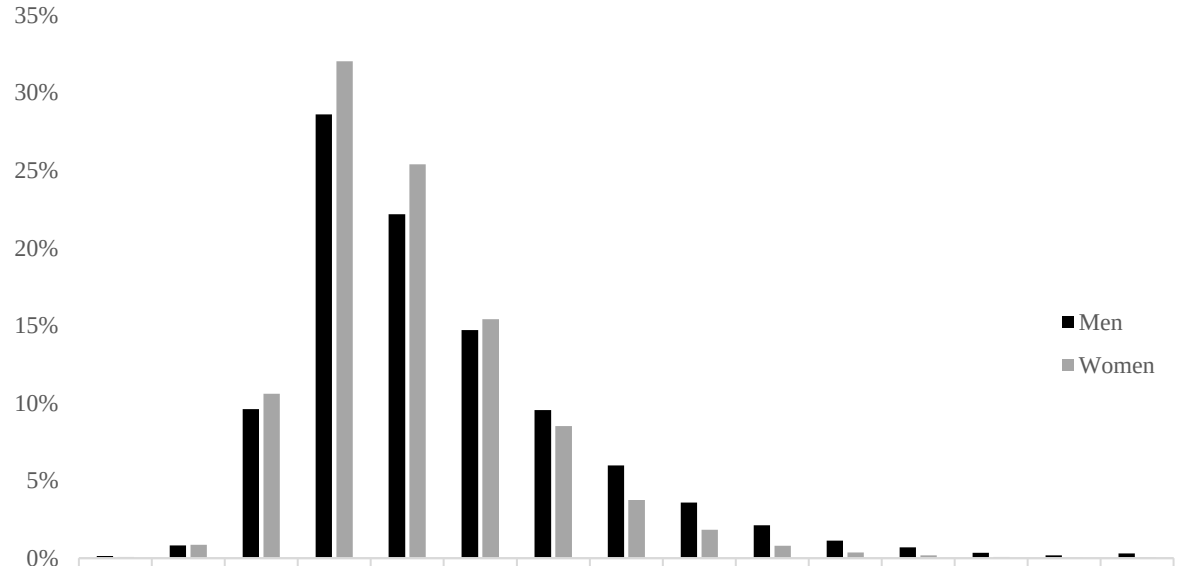
The data of this study has been derived from two big marathons in Germany in four consecutive years (Hamburg 2016-2018; Frankfurt 2015-2018). We selected these races because of the good data availability, the flat and even courses and furthermore to include races in both spring (Hamburg) and autumn (Frankfurt). Both races have large, heterogeneous athlete fields. Figure 1

[5] In the following, we refer to "year 1" as the year in which we measure the slowdown and in which the event of negative feedback occurs, and "year 2" as the period until the end of the following calendar year.

provides an overview of the distribution of slowdowns on the second half of the race in the initial sample.

Figure 1.

Distribution of slowdown among runners of the initial sample



	< 0,90	0,90-0,94	0,95-0,99	1,00-1,04	1,05-1,09	1,10-1,14	1,15-1,19	1,20-1,24	1,25-1,29	1,30-1,34	1,35-1,39	1,40-1,44	1,45-1,49	1,50-1,54	> 1,54
Men	0.1%	0.8%	9.6%	28.6%	22.2%	14.7%	9.6%	6.0%	3.6%	2.1%	1.1%	0.7%	0.3%	0.2%	0.3%
Women	0.1%	0.9%	10.6%	32.0%	25.4%	15.4%	8.5%	3.7%	1.8%	0.8%	0.4%	0.2%	0.1%	0.0%	0.1%

Legend: initial sample sample of 77.878 from 7 marathon races

On average, the athletes of our initial sample display a slowdown of 9.3% in the second half of the race (m=9.8%, f=7.7%)[6]. Looking at the professional athletes in particular indicates that an excessively fast initial race pace or a high variance in race pace do not represent an optimal race strategy. The 1% fastest athletes of each race and age group showed a deceleration of pace of 3.9% (m=3.9%, f=3.9%) in the second half of the race.

Of the initial sample of 78.778 athletes who crossed the finish line in the races under consideration, 3.206 athletes (4.1%) slowed down more than 30% in the second half of the race. Clear gender differences can be observed here: from the overall data pool 4.8% of the male

[6] The stronger slowdown among male athletes is in line with the results of numerous previous studies showing that men slow down more in the second half of the race, which is explained by higher tendency to overconfidence among men (see Krawczyk and Wilamowski (2017)).

athletes and 1.5% of the female athletes faced a slowdown of >30% in the second half of their race.

For our analyses, we focused on a sample of 800 athletes. Reason for this is the time-consuming manual collection and enrichment of the data. The 800 athletes are a stratified random sample. The strata are gender and nature of the feedback and are randomly drawn in equal parts. Thus, the sample is composed of 400 athletes with a slowdown of <30% (Group "negative feedback") and 400 athletes with a constant race and with a maximum speed deviation of 5% on the second half of the race (Group "no negative feedback"). Each subsample also has an equal share of 50% women and 50% men. Table 1 provides an overview of the general characteristics of our sample.

Table 1.

Overview sample characteristics

		Frequency	Proportion (%)
Gender	Female	400	50.0
	Male	400	50.0
Nature of feedback	negative	400	50.0
	not negative	400	50.0
First-Timer (first marathon)	no	493	61.6
	yes	307	38.4
Age Group	Age Group 1 | < 30 years	164	20.5
	Age Group 2 | 30 - 34 years	116	14.5
	Age Group 3 | 35 - 39 years	122	15.3
	Age Group 4 | 40 - 44 years	112	14.0
	Age Group 5 | 45 - 49 years	109	13.6
	Age Group 6 | 50 - 54 years	101	12.6
	Age Group 7 | ≥ 55 years	76	9.5
Race	Frankfurt 2016	107	13.4
	Frankfurt 2017	141	17.6
	Frankfurt 2018	145	18.1
	Hamburg 2016	83	10.4
	Hamburg 2017	113	14.1
	Hamburg 2018	124	15.5

Table 2 presents the racing behavior of the athletes in the analyzed year 1 separated by gender and nature of feedback. Women need on average 0.38h longer to finish the 42.2km. Athletes of both genders who experienced a strong slowdown in the race exhibit at the same time a slower overall race time. This difference is 1.1h for women and 1h for men and once again highlights the serious effect of a too fast initial pace and the resulting slowdown.

Table 2.

Descriptive statistics: behavior in year 1 by gender and nature of feedback

	Overall sample (n=800)	Negative feedback		No negative feedback	
		Women (n=200)	Men (n=200)	Women (n=200)	Men (n=200)
Overall net race time year 1 (in h)	4.39 (0.78)	5.15 (0.50)	4.70 (0.59)	4.01 (0.47)	3.71 (0.58)
Slowdown year 1	1.19 (0.19)	1.36 (0.07)	1.39 (0.09)	1.01 (0.02)	1.01 (0.02)

Legend: Mean (standard deviation)

The publicly accessible Internet database "marathon-ergebnis.de" was used to manually enrich the data set with supplementary information. "Marathon-ergebnis.de" contains results of all German marathon races of the last 10 years. The database covers not only large marathons, but also very small local races. With the help of the database, we were able to reconstruct whether the analyzed athletes participated in a marathon in Germany in the following year[7]. This matching was carried out manually, as we checked the plausibility of the results for each subject. In addition, we only took into account athletes who were clearly identifiable.[8] Furthermore, the database was used to determine whether the studied marathon was the athlete's first marathon ("first-timer") or whether the athlete had already gained prior marathon experience. In a second step, for all athletes who entered a marathon in year 2, the behavior of the athletes and in particular the overall running time and slowdown in year were traced in the (online available) result lists of the respective races.

[7] It can be assumed that there is no strong gender effect on the participation of German athletes in foreign races. In 2019, the proportion of women among German starters in the international Boston Marathon was 27% while in Frankfurt it was 23%.

[8] Requirement: Athletes must be identifiable via the database marathon-ergebnis.de. Criteria for this are German nationality and unambiguous attributability via name, age and/or club (e.g. not „Michael Smith, 30 years, no club“).

A: Descriptive Evidence: Re-entering a marathon in year 2

Regardless of gender and the nature of the previous racing experience, the probability of returning to race in the following year is 39.63% (m= 38.25%; f=41%).

Table 3.

Descriptive Statistics: re-entering in year 2 by gender and nature of feedback

	Overall sample (n=800)	Negative feedback		No negative feedback	
		Women (n=200)	Men (n=200)	Women (n=200)	Men (n=200)
Re-enter a marathon (1=yes, 0=no)	0.40 (0.49)	0.40 (0.49)	0.25 (0.43)	0.43 (0.50)	0.52 (0.50)

Legend: Mean (standard deviation)

When looking at the re-enter probability of the athletes separately by gender and nature of feedback (Table 3), clear gender-specific differences emerge. For men, there is a considerable effect of negative feedback: male athletes who experienced implicit negative feedback decide to run a marathon in the following year at a rate of 25%, whereas those without negative feedback re-entered a marathon at a rate of 52%. For women, the picture is contrary and female athletes with and without negative feedback differ by only 3 percentage points.

B: Descriptive Evidence: Behavior in year 2

The second part of the analyses specifies the behavior of athletes who decided to race in the following year. Athletes who experienced negative feedback in the last race clearly adjusted their race tactics. In this group slowdown is reduced by 22 percentage points (m= 22%; f=21%) and overall race time by 0.47h (m= 0.37h; f=0.53h). However, notwithstanding this improvement, both slowdown and overall net time of these athletes remained below the performance of the comparison group. This might be explainable by the fact that athletes who experience such a severe slowdown do not just inadvertently choose a too fast initial pace, but have fundamental difficulties with their race tactics and assessment of their performance, which cannot be simply taught in a short period of time. The average changes in net time and slowdown from year 1 to year 2 do not indicate any gender related differences.

Table 4.

Descriptive Statistics: behavior in year 2 by gender and nature of feedback

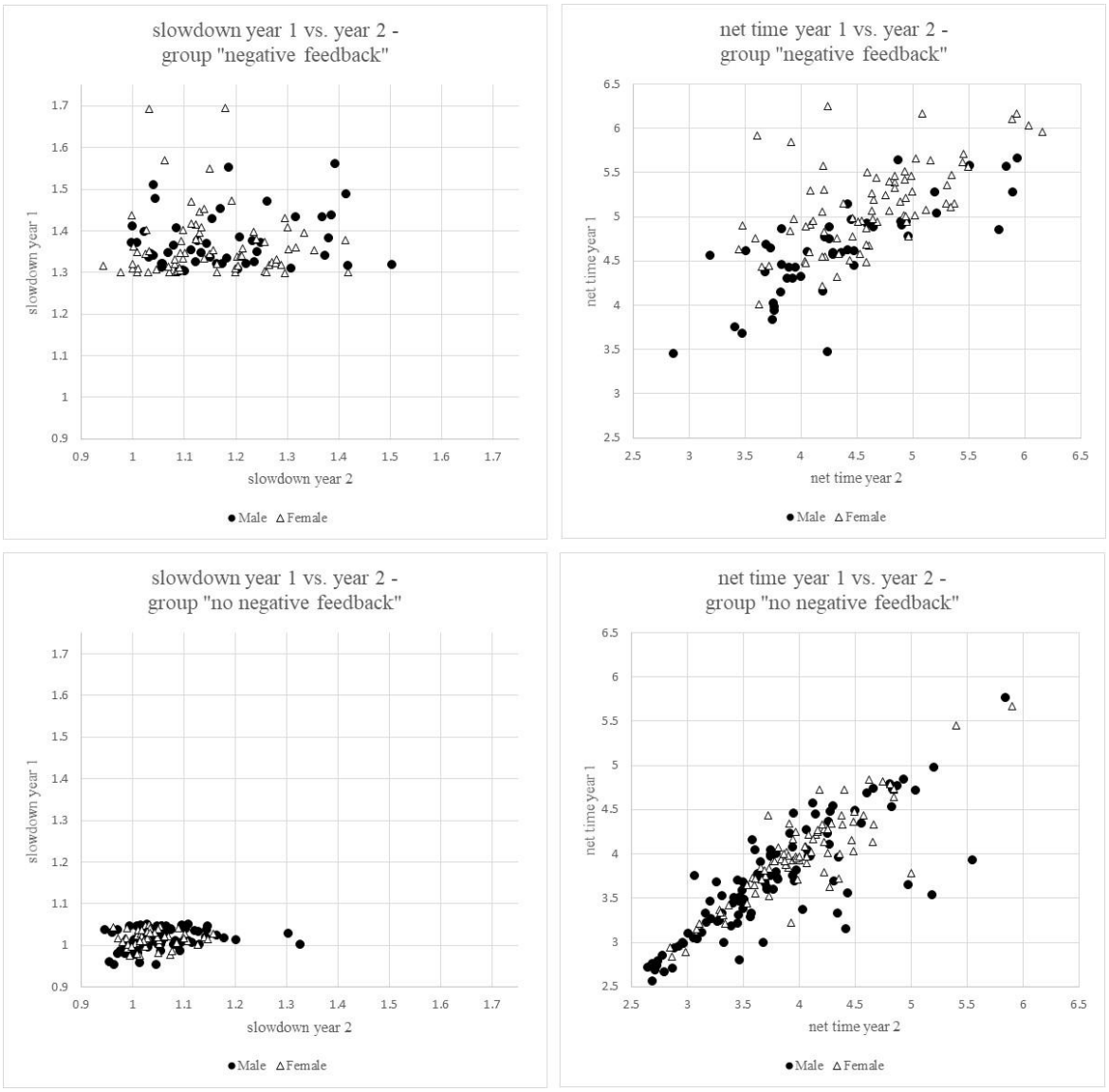
	Overall sample (n=317)	Negative feedback		No negative feedback	
		Women (n=79)	Men (n=49)	Women (n=85)	Men (n=104)
Overall net time year 2 (in h)	4.13 (0.71)	4.62 (0.60)	4.33 (0.71)	4.03 (0.53)	3.75 (0.68)
Slowdown year 2	1.10 (0.11)	1.15 (0.11)	1.17 (0.14)	1.04 (0.04)	1.05 (0.07)
Change net time $(\frac{\text{net time year 2}}{\text{net time year 1}})$	0.98 (0.11)	0.90 (0.09)	0.94 (0.10)	1.01 (0.07)	1.02 (0.11)
Change slowdown $(\frac{\text{slowdown year 2}}{\text{slowdown year 1}})$	0.95 (0.12)	0.85 (0.09)	0.85 (0.10)	1.03 (0.05)	1.02 (0.06)

Legend: Mean (standard deviation)

Figure 2 graphically depicts the development of the slowdown and overall net time of all individual athletes racing in year 2. The graph visually confirms the previous findings and demonstrates in particular that the adaptations differ across the two groups and that athletes with negative feedback in year 1 deviate more clearly from their behavior in the following year. Further, initially no clear gender differences in behavioral adjustment become apparent.

Figure 2.

Comparing development of slowdown and net time of year 1 and year 2 of each individual athlete



4. Econometric Findings

All statistical analyses were performed using the StataIC 16.0 software. Normality was checked with the Kolmogorov-Smirnoff test and visual inspection using Q-Q plots and histograms. The effects of the factors were tested using several regression methods. Two-sided p-values with less than 0.05 were taken as statistically significant.

A: Re-entering a marathon in year 1

First, we conducted probit regressions to investigate which factors influenced the tendency of athletes to re-enter a race. Since the descriptive data suggests significant interactions between "gender" and "negative feedback", we created four dummy variables to better isolate the effects. We used the "male + negative feedback" group as a reference group. In addition, we controlled for the respective race and year ("race dummy"), for first-timers and the age of the athletes. To avoid multicollinearity, we controlled for the speed of the athletes with the created variable "speed compared to median". For this we put the net race time of the athletes in relation to the median race time of their comparison group[9], which was formed according to their gender and age as well as the race and year. Consequently, a value greater than 1 indicates athletes who ran slower than their comparison group in the analyzed race. Table 5 gives an overview of mean and standard deviation of this created variable in our sample.

Table 5.

Overview variable "speed compared to Median"

	Overall sample (n=800)	Negative feedback		No negative feedback	
		Women (n=200)	Men (n=200)	Women (n=200)	Men (n=200)
Speed compared to Median	1.09 (0.19)	1.21 (0.12)	1.24 (0.15)	0.95 (0.11)	0.97 (0.15)

Legend: Mean (standard deviation)

Table 6 reports the marginal effects of the Model 1 probit regression, which examines the relationship between the tested variables and the tendency to re-enter. We find significant differences for gender and nature of prior feedback. For male athletes, the nature of the prior feedback causes a highly significant effect ($p<0.01$). Men who received negative feedback in year 1 were 22 percentage points less likely to run a marathon again in the following year (compared to men who did not receive negative feedback). In contrast, women are less affected by the previous experience: women who received negative feedback were 14.5 percentage points more

[9] A total of 98 comparison groups have thereby been distinguished (7 races; 7 age groups; 2 genders). "Speed compared to median" = $\frac{\text{net race time individual athlete}}{\text{net race time comparison group}}$.

likely to re-enter a marathon than men with negative feedback. Women seem to be considerably more resilient in dealing with prior negative experience.

Table 6.

Probit Regression: Explaining re-enter in year 2 with gender and nature of prior feedback

	Model 1
Dependent Variable	Re-enter in year 2 (yes = 1)
Model	Probit Regression
Independent Variable	
Gender x nature of feedback	
(Male + negative feedback)	
Male + no negative feedback	**.2200** (.0583)***
Female + no negative feedback	**.1454** (.0592)*
Female + negative feedback	**.1652** (.0494)**
First-Timer	**-.2401** (.0346)***
Age-Group	
(Age Group 1 | < 30 years)	
Age Group 2 | 30 - 34 years	**-.0947** (.0566)
Age Group 3 | 35 - 39 years	**-.0500** (.0562)
Age Group 4 | 40 - 44 years	**.0509** (.0585)
Age Group 5 | 45 - 49 years	**-.0322** (.0055)
Age Group 6 | 50 - 54 years	**.0551** (.0614)
Age Group 7 | ≥ 55 years	**-.0611** (.0697)
Speed compared to median	**-.0254** (.1290)
Race dummy	YES
Observations	800
Pseudo R2	0.0892

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

In addition, we find a highly significant negative relationship between first-timer and the tendency to re-enter in the following year. Athletes who ran their first race in the race under consideration were significantly less likely to run another race in the following year. Age and relative speed of the athletes do not significantly affect the re-enter of the athletes.

Further post-hoc Wald tests show that the groups "women with negative feedback" and "women without negative feedback" do not differ in their tendency to re-enter. For a more detailed understanding of gender differences in the impact of prior experience, Table 7 provides the regression models for the respective subsamples.

Table 7.

Probit Regression: Explaining re-enter in year 2 with nature of prior feedback for women and men

	Model 2	**Model 3**
Dependent variable	Re-enter in year 2 (yes = 1)	Re-enter in year 2 (yes = 1)
Model	Probit Regression	Probit Regression
Sample	Women	Men
Independent Variable		
Negative Feedback	-**.0140** (.0765)	-**.2321** (.0677)**
First-Timer	-**.2999** (.0485)***	-**.1968** (.0489)***
Age-Group		
(Age Group 1 | < 30 years)		
Age Group 2 | 30 - 34 years	-**.1276** (.0778)	-**.0770** (.0819)
Age Group 3 | 35 - 39 years	-**.0915** (.0767)	-**.0266** (.0819)
Age Group 4 | 40 - 44 years	**.1107** (.0585)	**.0166** (.0824)
Age Group 5 | 45 - 49 years	**.0322** (.0822)	-**.0816** (.0854)
Age Group 6 | 50 - 54 years	**.0111** (.0798)	**.0877** (.0876)
Age Group 7 | ≥ 55 years	**.0250** (.0871)	-**.1033** (.0885)
Speed compared to median	**.1238** (.2229)	-**.0282** (.1568)
Race dummy	YES	YES
Observations	400	400
Pseudo R2	0.0933	0.1163

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

Models 2 and 3 clearly corroborate the previous findings: In the male subsample, negative feedback significantly affects the re-entering of athletes. Male athletes who experienced a slowdown of more than 30% in year 1 are 23 percentage points less likely to re-enter in a marathon in the following year. In contrast, the nature of the preceding experience has no

significant effect on women's behavior or on their decision to participate in a marathon in the following year. Women appear to be resilient to a negative experience and unhindered by it face the competitive situation of a marathon again in the following year.

As in Model 1, first-timers in both subsamples tend to re-enter at a significantly lower rate in the following year. Table 8 provides more detailed results for separate subsamples of first-timers and no-first-timers. In Model 4, we confirm for the subsample of "No First-Timers", the significant effect of negative feedback among males as well as a significant gender difference. In contrast, for athletes competing in their first marathon (Model 5), we find no behavioral differences between men and women and only a slight effect of negative feedback for men.

These results further strenghten our assumptions by indicating that the gender differences in re-entreing are not caused by athletes who decided already in advance to run a marathon only as a "once in a lifetime" project but the findings are particularly true for athletes who have decided to run a marathon more than once in the past.

Table 8.

Probit Regression: Explaining re-enter in year 2 with nature of prior feedback for first-timers and non first-timers

	Model 4	**Model 5**
Dependent Variable	Re-enter in year 2 (yes = 1)	Re-enter in year 2 (yes = 1)
Model	Probit Regression	Probit Regression
Sample	No First-Timer	First-Timer
Independent Variable		
Gender x nature of feedback		
(Male + negative feedback)		
Male + no negative feedback	**.2441** (.0787)***	**.1872** (.0878)*
Female + no negative feedback	**.1859** (.0813)*	**.0809** (.0844)
Female + negative feedback	**.2582** (.0684)***	**.0550** (.0669)
Age-Group		
(Age Group 1 | < 30 years)		
Age Group 2 | 30 - 34 years	**-.1378** (.0823)	**-.0671** (.0711)
Age Group 3 | 35 - 39 years	**-.0893** (.0818)	**-.0183** (.0727)
Age Group 4 | 40 - 44 years	**-.0127** (.0835)	**.1474** (.0833)
Age Group 5 | 45 - 49 years	**-.0537** (.0798)	**-.0209** (.0867)
Age Group 6 | 50 - 54 years	**.0729** (.0799)	**-.0270** (.0904)
Age Group 7 | ≥ 55 years	**-.0915** (.0849)	**-.0461** (.1127)
Speed compared to median	**.0433** (.1701)	**-.1756** (.1975)
Race dummy	YES	YES
Observations	493	307
Pseudo R2	0.0515	0.0656

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

In a further step, the robustness of the results was tested using nearest neighbor matching. Table 9 reports the corresponding average treatment effects for the entire sample (Model 6) as well as for the subsamples of women (Model 7) and men (Model 8). The matching method again confirms our previous observations. In general, negative feedback has a negative effect on the probability of re-entering the race in the following year (Model 6). Athletes who slowdown more

than 30% in the race are significantly less likely to decide to run a marathon again the next year. Looking at the subsamples confirms that the impact of negative feedback is exclusively seen in male athletes (Model 8). Women are not influenced by the negative experience and irrespective of it decide for a re-enter (Model 7).

Table 9.

Nearest-Neighbor Matching: Effect of negative feedback on re-entering

		Model 6	**Model 7**	**Model 8**
	Dependent Variable	Re-enter	Re-enter	Re-enter
	Sample	All obs.	Women	Men
Average treatment effect (ATE)	Negative feedback	**-.1475** (.0568)**	**-.0650** (.0855)	**-.2475** (.0769)**
Matching variables	(Gender), Age Group, Race dummy, Speed compared to Median			
Observations		800	400	400

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

B: Behavior in Year 2

In the second part of the analyses, we study the behavior of the athletes in the following year (year 2). The sample sizes are reduced accordingly to the number of athletes who opted for a re-enter. Information about the athletes' behavior in the following year was gathered manually from the corresponding results lists that are publicly available online. However, for athletes who chose small, regional races sometimes no (split) times were available, which led to a further slight reduction of the sample size.

First we analyzed the change of slowdown and overall race time from year 1 to year 2 for each individual athlete. We not only considered the relative change, but also formed a binary dummy variable that marked an improvement in the following year with "1". We studied the general effect of negative feedback on behavior in the following year using nearest-neighbor matching. In all models, we found a significant difference in the group of athletes with negative feedback

(Table 10). Athletes who experienced negative feedback in the previous year improved significantly more both their race tactics (reduction of slowdown) as well as their overall net time.

Table 10.

Nearest neighbor matching: Behavior in year 2

		Model 14	**Model 15**	**Model 16**	**Model 17**
	Dependent Variable	slowdown y2 /slowdown y1	slowdown y2 < slowdown y1	net time y2 /net time y1	net time y2 < net time y1
Average treatment effect (ATE)	Negative feedback	**-.1983** (.0122)***	**.6828** (.0520)***	**-.0667** (.0171)***	**.3544** (.0723)***
Matching variables	Gender, Age Group, Race dummy, Speed compared to Median				
Observations		268	268	316	316

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

The following analyses in Table 11 and Table 12 outline the results of the nearest-neighbor matching for the two subsamples of female and male athletes. In contrast to the decision to restart, no gender differences in the adaptation of the behavior in the following year can be identified.

Regardless of gender, athletes who received negative feedback were more likely to reduce their slowdown and improved their overall net run time.

Table 11.

Nearest neighbor matching: Behavior in year 2 in subsample female athletes

		Model 18	**Model 19**	**Model 20**	**Model 21**
	Dependent Variable	slowdown y2 /slowdown y1	slowdown y2 < slowdown y1	net time y2 /net time y1	net time y2 < net time y1
Average treatment effect (ATE)	Negative feedback	**-.2043** **(**.0174)***	**.7404** (.0821)***	**-.0931** (.0119)***	**.4171** (.0791)***
Matching variables	Age Group, Race dummy, Speed compared to Median				
Observations		131	131	163	163

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

Table 12.

Nearest neighbor matching: Behavior in year 2 in subsample male athletes

		Model 22	**Model 23**	**Model 24**	**Model 25**
	Dependent Variable	slowdown y2 /slowdown y1	slowdown y2 < slowdown y1	net time y2 /net time y1	net time y2 < net time y1
Average treatment effect (ATE)	Negative feedback	**-.1919** (.0168)***	**.6569** (.0682)***	**-.0390** (.0310)	**.2941** (.1201)*
Matching variables	Age Group, Race dummy, Speed compared to Median				
Observations		137	137	153	153

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses.*

The significant difference in athletes with negative feedback can be attributed to the fact that our definition of negative feedback was pre-defined a slowdown of more than 30%. This translates into a potentially larger room for improvement for athletes who experienced a negative experience in the initial race. For this reason, we further split the analysis into subsamples according to the nature of their preceding feedback.

Once again, the previous findings could be confirmed: in all subsamples, we repeatedly found no gender-specific differences in the adaptation of race tactics in the form of a reduction of the slowdown or an improvement of the overall race time (Table 13&14). Women and men adapt their behavior to the same extent in the following year regardless of the nature of the previous feedback. Only first-timers show a slightly significant effect and athletes who competed in their first race in the previous year are more likely to run a faster time in the following year. Age or relative speed of the athletes do not explain the behavior in the following year.

Table 13.

Regression models: behavior in year 2 for subsample "negative feedback"

	Model 26	**Model 27**	**Model 28**	**Model 29**
Sub sample	Group "negative feedback"	Group "negative feedback"	Group "negative feedback"	Group "negative feedback"
Dependent Variable	slowdown y2/slowdown y1	slowdown y2 < slowdown y1	net time y2/net time y1	net time y2 < net time y1
Model	Linear Regression	Probit Regression	Linear Regression	Probit Regression
Independent Variable				
Female	**-.0155** (.0211)	**.0046** (.0637)	**-.0368** (.0191)	**.1006** (.1016)
First-Timer	**-.0275** (.0346)	**.0418** (.0606)	**-.0192** (.0191)	**.2255** (.0734)**
Age-Group				
(Age Group 1 | < 30 years)				
Age Group 2 | 30 - 34 years	**-.0219** (.0322)	**.0637** (.0647)	**.0500** (.0293)	**-.0770** (.1332)
Age Group 3 | 35 - 39 years	**-.0223** (.0301)	*omitted*	**-.0051** (.0276)	**-.0066** (.1220)
Age Group 4 | 40 - 44 years	**.0378** (.0294)	*omitted*	**.0336** (.0268)	**-.0742** (.1479)
Age Group 5 | 45 - 49 years	**.0271** (.0312)	**.0257** (.0829)	**.0370** (.0281)	**-.0348** (.1207)
Age Group 6 | 50 - 54 years	**.0150** (.0317)	**-.0115** (.1044)	**.0157** (.0280)	*omitted*
Age Group 7 | ≥ 55 years	**-.0026** (.0318)	*omitted*	**.0301** (.0288)	**-.0852** (.1441)
Speed compared to median	**.2141** (.0753)**	**-.5940** (.2881)*	**-.1366** (.0678)*	**-.2061** (.3197)
Constant	**.5827** (.0964)***	-	**1.0792** (.0866)***	-
Race dummy	YES	YES	YES	YES
Observations	121	75	128	111
Adj R-squared / Pseudo R2	0.0685	0.1660	0.1605	0.2491

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses. For probit models marginal effects are depicted.*

Table 14.

Regression models: behavior in year 2 for subsample "no negative feedback"

	Model 30	**Model 31**	**Model 32**	**Model 33**
Sub sample	Group "no negative feedback"	Group "no negative feedback"	Group "no negative feedback"	Group "no negative feedback"
Dependent Variable	slowdown y2/slowdown y1	slowdown y2 < slowdown y1	net time y2/net time y1	net time y2 < net time y1
Model	Linear Regression	Probit Regression	Linear Regression	Probit Regression
Independent Variable				
Female	**-.0072** (.0100)	**.0470** (.0810)	**-.0136** (.0141)	**.0724** (.0742)
First-Timer	**-.0032** (.0346)	**.0806** (.0968)	**-.0360** (.0169)*	**.0854** (.0897)
Age-Group				
(Age Group 1 | < 30 years)				
Age Group 2 | 30 - 34 years	**-.0051** (.0199)	**-.0559** (.0566)	**.0222** (.0280)	**.0221** (.1437)
Age Group 3 | 35 - 39 years	**.0026** (.0176)	**-.0310** (.1468)	**.0618** (.0245)*	**-.0689** (.1291)
Age Group 4 | 40 - 44 years	**-.0087** (.0169)	**-.0407** (.1399)	**.0539** (.0239)*	**-.2537** (.2151)*
Age Group 5 | 45 - 49 years	**.0074** (.0174)	**-.2127** (.1313)	**.0138** (.0250)	**-.1594** (.1327)
Age Group 6 | 50 - 54 years	**-.0095** (.0165)	**-.0975** (.1348)	**.0347** (.0237)	**-.2472** (.1245)*
Age Group 7 | ≥ 55 years	**-.0003** (.0211)	**-.0975** (.1348)	**.0112** (.0298)	**-.3215** (.1531)*
Speed compared to median	**.0664** (.0342)	**-.0608** (.1765)	**-.0716** (.0496)	**.1188** (.1290)
Constant	**.9627** (.0356)***	-	**1.0756** (.0509)***	-
Race dummy	YES	YES	YES	YES
Observations	147	147	188	188
Adj R-squared / Pseudo R2	0.0681	0.0740	0.0439	0.0587

Legend: **** denotes significance <1% ** denotes significance at 1%. * denotes significance at 5%. Robust standard errors are in parentheses. For probit models marginal effects are depicted.*

5. Summary and Implications

The presented results confirm previous research and clearly show that feedback affects behavior of individuals in repeated competitions and that negative feedback in particular reduces the likelihood of re-participation.

However, in contrast to prior studies, our results indicate that negative feedback only affects the behavior of men. Studying the behavior of amateur marathon runners we find that negative feedback causes men to drop out of competitive situations while women seem to be resilient and unaffected by the nature of prior feedback. These results contrast with earlier findings showing that women responded to loss by reducing their performance (Buser, 2016).

The reasons for the contrasting results lie most likely in the advantage of the real-life context of our study. Unlike in a laboratory design, where incentives are clearly formulated and there is a distinct objective of the designed game as well as the definition of success and failure, amateur marathon decide to participate for a wide variety of reasons. These reasons range from sense of personal achievment (and associated status) of a marathon finish, to reaching new personal best times, to competitive comparison with other contestants. The underlying motivation clearly shapes how negative feedback affects the runner. For example, athletes who participate in a marathon purely for intrinsic enjoyment might be less affected by negative feedback. Prior findings suggests that women and men expose themselves to competitive conditions for different reasons and strongly differ in their preferences when facing competitive situations (Niederle, 2016). The different approaches and attitudes towards the given circumstances could be the reason for the resilient reaction of female athletes to the negative feedback.

Furthermore, among athletes who chose to re-enter, no gender difference in change of behavior could be seen in the following year. Both women and men improved the assessment of their abilities in the following year and ran a more constant race. Our results confirm not only that gender gap in overconfidence dissolves in repeated competition, but also that this phenomenon also occurs when individuals are confronted with negative feedback.

Participation in a marathon is in many respects very similar to the conditions on the labor market: in both, participants repeatedly experience and respond to successes and losses; in both,

repeatedly decisions on whether and how to (re)enter a competitive situation are necessary; in both participants voluntarily decide on the type of participation and, in particular, with what intensity or underlying motivation and competitiveness they participate.

The results of the present study particularly help to understand the existing gender gap in the competitive field of the labor market and help to design possible interventions and structured measures. The insights are obtained by comparing the circumstances and conditions in sports competitions (where the gender gap dissolves in repeated competition) with the professional world (where the gender gap persists in repeated competition) and analyzing what institutional factors lead to the different outcomes. By combining the results of the present study and findings from previous research, three potential fields of action manifest.

5.1. Systematic feedback: highly-frequented, cost-free, unambiguous, and mandatory

The present study confirms a large number of previous research findings: feedback causes gender differences in competitivity and self-confidence to disappear. In marathon running women are constantly exposed to implicit feedback: on the one hand when comparing the actual pace with adjacent runners, on the other hand as a physiological reaction of the own body to the chosen and/or desired pace. In a professional context, highly frequented feedback could as well help to prevent recurring phases of insecurity of women. Particularly in the labor market, objective feedback on one's own performance is often incomplete and above all not cost-free. We know from adjacent research that women have a clear aversion to feedback and are less willing to buy feedback if it is not free or obligatory (Wozniak et al., 2016). For example, students entering the labor market do not have transparent information on their relative position and strengths. This is one possible explanation why women with equal educational levels enter the labor market in less attractive jobs. Several previous studies showed for example that despite equal performance female students are significantly less confident about their educational abilities (Blanch et al., 2008; Michie et al., 2001). Dickerson and Taylor (2000) showed that women only applied for a job if they matched all listed requirements, whereas men applied regardless of missing skills. Providing systematic, complimentary assistance to female students at the end of their studies, e.g. with experienced career mentors, could help women to get a more realistic and objective assessment of their own strengths and potentially apply for more suitable and/or challenging jobs.

Furthermore, our results indicate that feedback for employees in competitive environments should be unambiguous, honest, and even negative. In our study, negative feedback caused low-ability men to drop out of repeated competition, which reduced part of the gender gap.

5.2. Transparent "tournament conditions”

An important difference in the sports context is the clarity and transparency of the competition conditions. Athletes can precisely calculate in advance which pace is necessary for a 3:00 h marathon or that for a finish on the podium the third fastest time in the respective category is required. This differs considerably from the transparency of tournament conditions in the workplace: Very often there are incomplete information what performance and achievements are required for a promotion or what additional criteria, apart from performance, are decisive for a salary increase. Transparent programs such as automatic promotions based on performance might help women particularly to strengthen their self-confidence and remain competitive. This has already been demonstrated by field experiments in the context of sales. In sales, there is generally a very high transparency of conditions, clear objectives, and monetary consequences. A study in sales organizations revealed no gender differences in performance in this transparent environment and indicated that women had a lower propensity to leave this competitive environment (Moncrief et al., 2000).

5.3. Empowerment of women in early stages of their career

The present study clearly confirms that women who selected themselves into a competitive environment are particularly good at coping with negative experiences and remain in competition regardless of any difficulties that may arise (see also John, 2017). Systematic interventions should address this by supporting women in the early stages of their careers and helping them overcome the first barriers until they have established themselves in the competitive environment. Particular focus should be given to women in high performance environments, as there is a high incidence and necessity to undertake high-risk decisions. For example, early, intensive mentoring of high-performing female students could encourage them to enter attractive and competitive jobs. An enhanced encouragement of women in low and middle management also results in an increase in the total pool and variety of female candidates for senior management positions. Additionally, an increased proportion of women in every level leads to a positive upward spiral and to the

strengthening of the female workforce: when women are underrepresented compared to men, they achieve lower performance and exit the competition at an earlier stage (Gneezy et al., 2003; Hogarth et al., 2012).

The present study holds a few limitations, but these primarily offer opportunities for further research. On the one hand, in our study we focus exclusively on participation in marathon races in the following year. Other studies could further investigate whether male athletes withdraw completely from racing after a setback, or possibly switch to other distances such as half marathon and ultra marathon or alternative sports such as triathlon. Furthermore, in addition to psychological attributes, physiological factors also play a role in sport such as women's skeletal muscle showing less fatigability or gender differences in susceptibility to muscle glycogen depletion (see e.g. Hunter, 2014; Roepstorff et al., 2002). Possibly the same slowdown may feel subjectively less painful for women than for men, which is why men react more strongly to negative feedback in sport. Therefore, further studies are needed to understand how gender differences in the response to negative feedback occur in other natural environments where physiological differences play a negligible role. Finally, our study investigates the effect of implicit, relative feedback. Implicit feedback entails that the feedback recipient must independently receive, process, assess and interpret the information. The focus of subsequent studies should be to understand to what extent the suggested measures can be transferred to explicit feedback and what adjustments to existing interventions are necessary to reduce the gender gap in the labor market systematically.

References

Abbiss, C. R., & Laursen, P. B. (2008). Describing and understanding pacing strategies during athletic competition. *Sports Medicine*, *38*(3), 239–252. https://doi.org/10.2165/00007256-200838030-00004

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., & Tungodden, B. (2016). What explains the gender gap in college track dropout? Experimental and administrative evidence. *American Economic Review, 106*(5), 296–302. https://doi.org/10.1257/aer.p20161075

Antonovics, K., Arcidiacono, P., & Walsh, R. (2009). The effects of gender interactions in the lab and in the field. *Review of Economics and Statistics*, *91*(1), 152–162. https://doi.org/10.1162/rest.91.1.152

Atkinson, S. M., Baird, S. B., & Frye, M. B. (2003). Do female mutual fund managers manage differently? *Journal of Financial Research*, *26*(1), 1–18. https://doi.org/10.1111/1475-6803.00041

Azmat, G., & Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, *94*(7-8), 435–452. https://doi.org/10.1016/j.jpubeco.2010.04.001

Balafoutas, L., & Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science, 335*(6068), 579–582. https://doi.org/10.1126/science.1211180

Balafoutas, L., & Sutter, M. (2019). How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior. *European Economic Review*, *118*, 1–13. https://doi.org/10.1016/j.euroecorev.2019.05.005

Bélanger, J. J., Lafrenière, M.-A. K., Vallerand, R. J., & Kruglanski, A. W. (2013). Driven by fear: The effect of success and failure information on passionate individuals' performance. *Journal of Personality and Social Psychology*, *104*(1), 180–195. https://doi.org/10.1037/a0029585

Bench, S. W., Lench, H. C., Liew, J., Miner, K., & Flores, S. A. (2015). Gender gaps in overestimation of math performance. *Sex Roles*, *72*(11), 536–546. https://doi.org/10.1007/s11199-015-0486-9

Berlin, N., & Dargnies, M.-P. (2016). Gender differences in reactions to feedback and willingness to compete. *Journal of Economic Behavior & Organization*, *130*, 320–336. https://doi.org/10.1016/j.jebo.2016.08.002

Beyer, S., & Bowden, E. M. (1997). Gender differences in self-perceptions: Convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin*, *23*(2), 157–172. https://doi.org/10.1177/0146167297232005

Black, S. E., & Strahan, P. E. (2001). The division of spoils: Rent-sharing and discrimination in a regulated industry. *American Economic Review*, *91*(4), 814–831. https://doi.org/10.1257/aer.91.4.814

Blanch, D. C., Hall, J. A., Roter, D. L., & Frankel, R. M. (2008). Medical student gender and issues of confidence. *Patient Education and Counseling*, *72*(3), 374–381. https://doi.org/10.1016/j.pec.2008.05.021

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789–865. https://doi.org/10.1257/jel.20160995

Bognanno, M. L. (2001). Corporate tournaments. *Journal of Labor Economics*, *19*(2), 290–315. https://doi.org/10.1086/319562

Boll, C., & Leppin, J. S. (2015). Die geschlechtsspezifische Lohnlücke in Deutschland: Umfang, Ursachen und Interpretation. *Wirtschaftsdienst*, *95*(4), 249–254. https://doi.org/10.1007/s10273-015-1814-y

Buser, T. (2016). The impact of losing in a competition on the willingness to seek further challenges. *Management Science*, *62*(12), 3439–3449. https://doi.org/10.1287/mnsc.2015.2321

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409–1447. https://doi.org/10.3386/w18576

Card, D., Cardoso, A. R., & Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, *131*(2), 633–686. https://doi.org/10.1093/qje/qjv038

Clark, A., & Oswald, A. (1996). Satisfaction and comparison income. *Journal of Public Economics*, *61*(3), 359–381. https://doi.org/10.1016/0047-2727(95)01564-7

Cotton, C., McIntyre, F., & Price, J. (2013). Gender differences in repeated competition: Evidence from school math contests. *Journal of Economic Behavior & Organization*, *86*, 52–66. https://doi.org/10.1016/j.jebo.2012.12.029

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–474. https://doi.org/10.1257/jel.47.2.448

Deaner, R. O., Carter, R. E., Joyner, M. J., & Hunter, S. K. (2015). Men are more likely than women to slow in the marathon. *Medicine and Science in Sports and Exercise*, *47*(3), 607–616. https://doi.org/10.1249/MSS.0000000000000432

Dickerson, A., & Taylor, M. A. (2000). Self-limiting behavior in women. *Group & Organization Management*, *25*(2), 191–210. https://doi.org/10.1177/1059601100252006

Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*, *101*(2), 556–590. https://doi.org/10.1257/aer.101.2.556

Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: Ii. The contingencies of evaluative feedback in the classroom and iii. An experimental analysis. *Developmental Psychology*, *14*(3), 268–276. https://doi.org/10.1037/0012-1649.14.3.268

Easterlin, E. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization*, *27*(1), 35–47. https://doi.org/10.1016/0167-2681(95)00003-B

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology*, *84*(1), 5–17.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*(1), 218–232.

Elliot, A. J., & Thrash, T. M. (2004). The intergenerational transmission of fear of failure. *Personality and Social Psychology Bulletin*, *30*(8), 957–971. https://doi.org/10.1177/0146167203262024

Eriksson, T. (1999). Executive compensation and tournament theory: Empirical tests on danish data. *Journal of Labor Economics*, *17*(2), 262–280. https://doi.org/10.1086/209920

Ertac, S., & Szentes, B. (2011). The effect of information on gender differences in competitiveness: Experimental evidence. *Koç University-TUSIAD Economic Research Forum Working Papers*(1104).

EWOB. (2019). *Gender Diversity Index 2019*. Brussels. European Women on Boards.

Flory, J. A., Gneezy, U., Leonard, K. L., & List, J. A. (2018). Gender, age, and competition: A disappearing gap? *Journal of Economic Behavior & Organization*, *150*, 256–276. https://doi.org/10.1016/j.jebo.2018.03.027

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, *102*(1), 469–503. https://doi.org/10.1257/aer.102.1.469

Gill, D., & Prowse, V. (2014). Gender differences and dynamics in competition: The role of luck. *Quantitative Economics*, *5*(2), 351–376. https://doi.org/10.3982/QE309

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074. https://doi.org/10.1162/00335530360698496

Gneezy, U., & Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review*, *94*(2), 377–381. https://doi.org/10.1257/0002828041301821

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091–1119. https://doi.org/10.1257/aer.104.4.1091

Günther, C., Ekinci, N. A., Schwieren, C., & Strobel, M. (2010). Women can't jump? - an experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior & Organization*, *75*(3), 395–401. https://doi.org/10.1016/j.jebo.2010.05.003

Haenni, S. (2019). Ever tried. Ever failed. No matter? On the demotivational effect of losing in repeated competitions. *Games and Economic Behavior*, *115*, 346–362. https://doi.org/10.1016/j.geb.2019.03.012

Hogarth, R. M., Karelaia, N., & Trujillo, C. A. (2012). When should I quit? Gender differences in exiting competitions. *Journal of Economic Behavior & Organization*, *83*(1), 136–150. https://doi.org/10.1016/j.jebo.2011.06.021

Hunter, S. K. (2014). Sex differences in human fatigability: Mechanisms and insight to physiological responses. *Acta Physiologica (Oxford, England)*, *210*(4), 768–789. https://doi.org/10.1111/apha.12234

John, J. P. (2017). Gender differences and the effect of facing harder competition. *Journal of Economic Behavior & Organization*, *143*, 201–222. https://doi.org/10.1016/j.jebo.2017.08.012

Kamas, L., & Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, *83*(1), 82–97. https://doi.org/10.1016/j.jebo.2011.06.013

Kimball, M. M., & Gray, V. A. (1982). Feedback and performance expectancies in an academic setting. *Sex Roles, 8*(9), 999–1007. https://doi.org/10.1007/BF00290024

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Krawczyk, M., & Wilamowski, M. (2017). Are we all overconfident in the long run? Evidence from one million marathon participants. *Journal of Behavioral Decision Making, 30*(3), 719–730. https://doi.org/10.1002/bdm.1984

Kuhnen, C. M., & Tymula, A. (2012). Feedback, self-esteem, and performance in organizations. *Management Science, 58*(1), 94–113.

Legge, S., & Schmid, L. (2013). Rankings, success, and individual performance: Evidence from a natural experiment. *University of St. Gallen Law & Economics, Working Paper No. 2013-13.* https://doi.org/10.2139/ssrn.2161903

Lenney, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin, 84*(1), 1–13. https://doi.org/10.1037/0033-2909.84.1.1

Lundeberg, M. A., Fox, P. W., & Puncochar, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of Educational Psychology, 86*(1), 114–121. https://doi.org/10.1037/0022-0663.86.1.114

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics, 120*(3), 963–1002. https://doi.org/10.1093/qje/120.3.963

March, D. S., Vanderburgh, P. M., Titlebaum, P. J., & Hoops, M. L. (2011). Age, sex, and finish time as determinants of pacing in the marathon. *Journal of Strength and Conditioning Research, 25*(2), 386–391. https://doi.org/10.1519/JSC.0b013e3181bffd0f

Masclet, D., Peterle, E., & Larribeau, S. (2015). Gender differences in tournament and flat-wage schemes: An experimental study. *Journal of Economic Psychology, 47,* 103–115. https://doi.org/10.1016/j.joep.2015.01.003

McCarty, P. A. (1986). Effects of feedback on the self-confidence of men and women. *Academy of Management Journal*, *29*(4), 840–847. https://doi.org/10.2307/255950

McGraw, A. P., Mellers, B. A., & Ritov, I. (2004). The affective costs of overconfidence. *Journal of Behavioral Decision Making*, *17*(4), 281–295. https://doi.org/10.1002/bdm.472

Michie, F., Glachan, M., & Bray, D. (2001). An evaluation of factors influencing the academic self-concept, self-esteem and academic stress for direct and re-entry students in higher education. *Educational Psychology*, *21*(4), 455–472. https://doi.org/10.1080/01443410120090830

Moncrief, W. C., Babakus, E., Cravens, D. W., & Johnston, M. W. (2000). Examining gender differences in field sales organizations. *Journal of Business Research*, *49*(3), 245–257. https://doi.org/10.1016/S0148-2963(99)00019-3

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101. https://doi.org/10.1162/qjec.122.3.1067

Niederle, M. (2016). Gender. In J. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (2nd ed., pp. 481–553). Princeton University Press.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1–16. https://doi.org/10.1287/mnsc.1120.1602

Price, J. (2008). Gender differences in the response to competition. *ILR Review*, *61*(3), 320–333. https://doi.org/10.1177/001979390806100303

Roepstorff, C., Steffensen, C. H., Madsen, M., Stallknecht, B., Kanstrup, I.-L., Richter, E. A., & Kiens, B. (2002). Gender differences in substrate utilization during submaximal exercise in endurance-trained subjects. *American Journal of Physiology. Endocrinology and Metabolism*, *282*(2), E435-47. https://doi.org/10.1152/ajpendo.00266.2001

Rosenqvist, O., & Nordström Skans, O. (2015). Confidence enhanced performance? – the causal effects of success on future performance in professional golf tournaments. *Journal of Economic Behavior & Organization*, *117*, 281–295. https://doi.org/10.1016/j.jebo.2015.06.020

Storek, J., & Furnham, A. (2014). Gender and task confidence as predictors of the domain-masculine intelligence type (dmiq). *Personality and Individual Differences*, *69*, 43–49. https://doi.org/10.1016/j.paid.2014.05.006

Sutter, M., & Glätzle-Rützler, D. (2010). *Gender Differences in Competition Emerge Early in Life* (Institute of Labor Economics (IZA) No. 5015).

Szymanski, K., & Harkins, S. G. (1987). Social loafing and self-evaluation with a social standard. *Journal of Personality and Social Psychology*, *53*(5), 891–897. https://doi.org/10.1037/0022-3514.53.5.891

White, P. H., Kjelgaard, M. M., & Harkins, S. G. (1995). Testing the contribution of self-evaluation to goal-setting effects. *Journal of Personality and Social Psychology*, *69*(1), 69–79. https://doi.org/10.1037/0022-3514.69.1.69

Wozniak, D. (2011). Gender differences in a market with relative performance feedback: Professional tennis players. *Journal of Economic Behavior & Organization*, *83*(1), 158–171. https://doi.org/10.1016/j.jebo.2011.06.020

Wozniak, D., Harbaugh, W. T., & Mayr, U. (2011). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, *32*(1), 161–198. https://doi.org/10.1086/673324

Wozniak, D., Harbaugh, W. T., & Mayr, U. (2016). *The Effect of Feedback on Gender Differences in Competitive Choices*.

Zillman, C. (2019). *The fortune 500 has more female CEOs than ever before.* Fortune 500. https://fortune.com/2019/05/16/fortune-500-female-ceos/