Lovász, Anna; Szabó-Morvai, Ágnes; Cukrowska-Torzewska, Ewa; Rigó, Mariann; Kiss, Andrea

Working Paper

One size fits all? Gender differences in the effect of subjective feedback

Budapest Working Papers on the Labour Market, No. BWP - 2017/5

Provided in Cooperation with:
Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences

Suggested Citation: Lovász, Anna; Szabó-Morvai, Ágnes; Cukrowska-Torzewska, Ewa; Rigó, Mariann; Kiss, Andrea (2017) : One size fits all? Gender differences in the effect of subjective feedback, Budapest Working Papers on the Labour Market, No. BWP - 2017/5, ISBN 978-615-5754-00-5, Hungarian Academy of Sciences, Institute of Economics, Centre for Economic and Regional Studies, Budapest

This Version is available at:
https://hdl.handle.net/10419/200356

BUDAPEST WORKING PAPERS
ON THE LABOUR MARKET
BWP – 2017/5

One Size Fits All? Gender Differences in the Effect of Subjective Feedback

ANNA LOVÁSZ, EWA CUKROWSKA-TORZEWSKA,
MARIANN RIGÓ, ÁGNES SZABÓ-MORVAI, ANDREA KISS

BWP 2017/5

INSTITUTE OF ECONOMICS, CENTRE FOR ECONOMIC AND REGIONAL STUDIES
HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, 2017

Budapest Working Papers on the Labour Market

BWP – 2017/5

Institute of Economics, Centre for Economic and Regional Studies,
Hungarian Academy of Sciences

One Size Fits All? Gender Differences in the Effect of Subjective Feedback

Authors:

Anna Lovász
research fellow
Institute of Economics
Centre for Economic and Regional Studies, Hungarian Academy of Sciences
and ELTE University
email: lovasz.anna@krtk.mta.hu

Ágnes Szabó-Morvai
research fellow
Institute of Economics
Centre for Economic and Regional Studies, Hungarian Academy of Sciences
and HETFA Institute
email: szabomorvai.agnes@krtk.mta.hu

Ewa Cukrowska-Torzewska
University of Warsaw, Faculty of Economic Sciences

Mariann Rigó
Institute of Gerontology at TU Dortmund University

Andrea Kiss
Duke University, North Carolina

June 2017

ISBN 978-615-5754-00-5
ISSN 1785 3788

One Size Fits All? Gender Differences in the Effect of Subjective Feedback

Anna Lovász, Ewa Cukrowska-Torzewska, Mariann Rigó,
Ágnes Szabó-Morvai, Andrea Kiss

Abstract

The effect of objective feedback on performance is often studied, while subjective feedback is largely neglected in the economics literature. We estimate the impact of positive subjective feedback - encouragement and praise - on effort and performance, and compare the effect by gender. We use a computer game, during which players are randomly chosen to be given either no feedback (control) or positive subjective feedback (treatment), and analyze the treatment effect on effort (clicks) and performance (score). Based on previous economic and psychology theories, we test the pathways through which subjective feedback can have an impact: on (1) effort, due to the updating of expected performance or direct (dis)utility from the feedback, or (2) marginal productivity. The results point to significant differences in the mean effects of subjective feedback by gender. For women, encouragement has a significant positive effect while praise has a significant negative effect on performance, while men are less responsive to subjective feedback in general. Gender differences are mostly explained by different confidence distributions, while there are no gender differences in treatment effects if confidence level is held fixed. The effects are mostly realized through changes in effort. These results suggest that better targeted supervisory communication in schools or workplaces can improve the performance of lower-confidence individuals and thereby decrease the gender gap in performance.

Keywords: gender differences, supervisory feedback, experimental economics

JEL codes: C90, D03, J16, M54

Acknowledgements:

The project leading to this application has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 691676. We would like to thank Andor Zoldesi for the development of the online game and research website. We are grateful to Daniel Horn, Hubert Janos Kiss, Barbara Pertold-Gebicka, and participants of the Education Economics Network workshops in Leuven and Maastricht, the Szirák Labor Market Conference, the RECENS research group, and seminars at our institutions for their comments. This research is being carried out through collaboration via the Virtual Research Collaboration on Gender and Family in the Labor Market.

Nemek közötti különbségek a szubjektív visszajelzés hatásában

Lovász Anna, Ewa Cukrowska-Torzewska, Rigó Mariann,
Szabó-Morvai Ágnes, Kiss Andrea

Összefoglaló

Az objektív visszajelzések teljesítményre gyakorolt hatása gyakran vizsgált kérdéskör, azonban a közgazdasági irodalomban a szubjektív visszajelzésekről nem állnak rendelkezésre kutatási eredmények. Kísérletünkben a szubjektív visszajelzések – a dicséret és bíztatás – teljesítményre és erőfeszítésre gyakorolt hatását becsüljük meg, majd a hatást nemenként hasonlítjuk össze. Egy online játékot használunk, amely során a játékosok véletlenszerűen kapnak vagy nem kapnak szubjektív visszajelzést. Megvizsgáljuk, ez hogyan hat az erőfeszítésre (kattintások száma), illetve a teljesítményükre (pontszám). Előző közgazdasági és pszichológiai irodalmak alapján azt is vizsgáljuk, hogy a hatás milyen csatornákon keresztül valósulhat meg: (1) az erőfeszítésen keresztül, vagy (2) határtermelékenységen keresztül. Az eredmények alapján szignifikáns különbségek vannak a nemek között a szubjektív visszajelzés hatásában. A nők teljesítményére átlagosan pozitívan hat a bíztatásés negatívan a dicséret, míg a férfiak kevésbé érzékenyek mindkét visszajelzés típusra. A nemek közötti átlagos különbségeket majdnem teljes egészben magyarázza a magabiztosságukban való eltérés. Azonos magabiztosságú csoportokon belül nincsen ststisztikailag értelmezhető eltérés a nemek között a szubjektív visszajelzés hatásában. Az eredmények arra utalnak, hogy az iskolai és munkahelyi kommunikáció célzottabbá tétele jelentősen javíthatja az alacsonyabb magabiztosságú egyének teljesítményét, és ezáltal csökkentheti a nemek közti különbségeket is.

JEL kódok: C90, D03, J16, M54

Kulcsszavak: nemek közötti különbségek, vezetői visszajelzések, kísérleti gazdaságtan

1. INTRODUCTION

Gender differences in various psychological traits – competitiveness, risk aversion, cooperation, altruism, self-confidence - have been documented in a significant new strand of laboratory and field experiments.[1] Such differences may contribute to existing gender gaps in both educational and labor market outcomes, as a few previous studies have shown using datasets with real-life outcome measures (Buser et al 2014, Ors et al 2013).[2] As Niederle (2016) discusses, a natural next question is what we could do about these differences. Two main research directions seek to provide an answer. The first focuses on determining which factors - such as hormones, age, socio-economic status, or culture – contribute to the gender gaps in preferences and traits, to see whether they can be altered (*"fixing women"*). The second focuses on how elements of our environment or market design could be altered to achieve outcomes that more closely reflect underlying abilities, instead of creating gender gaps because they favor certain psychological traits that are unequally distributed by gender (*"fixing institutions"*).[3] We add to this line of research by testing for gender differences in the effect of one specific element of the environment: subjective feedback, in the form of encouragement and praise. We argue that the way people receive feedback on their performance largely affects their subsequent performance, and that different types of feedback are suitable for different types of individuals.

We use a simple online game with randomized treatment to test whether individual performance is affected by such feedback, and how this effect varies by gender. The game is available on a website, which allows for data collection via social media outlets. Players are randomly chosen to be given either no subjective feedback or positive subjective feedback, in the form of simple phrases and emoticons appearing in textboxes that pop up during the game.[4] Detailed data on players' performance (score) and effort (clicks) is linked to pre-game survey information on basic demographics and task-specific self-confidence. This is used to compare outcomes by treatment type and gender, and to assess the roles of the effort and marginal productivity channels and self-confidence in the observed treatment effects. Despite the very brief interaction and limited treatment given during the game, our findings point to

[1] See Eckel and Grossman (2008), Croson and Gneezy (2009), and Niederle (2016) for reviews of the experimental evidence on gender differences.
[2] Bertrand (2011), and Azmat and Petrongolo (2014) assess the experimental findings on gender differences in traits and preferences, and their implications for labor economics.
[3] An example of such a policy that may decrease gender gaps in competitiveness is given, for example, by Ertac and Szentes (2010) and Wozniak et al (2014), who find that the provision of relative performance feedback during a task may eliminate the gender gap in the choice to compete.
[4] Encouragement and praise are given as separate treatments that differ in both their content (future vs. past performance), and timing (time-based or performance-based), with separate corresponding control groups for each, as discussed later on.

a significant gender difference which is driven by the confidence distribution by gender. This bears clear implications for HR and education policies: first, the performance of certain individuals and groups may be highly affected by supervisory communication. Second, individual as well as class and workforce level outcomes can be improved significantly if individuals receive communication that is better targeted to their personality.

Subjective feedback can be thought of as an element of the supervisory communication between teachers and students or managers and workers that is relatively easy to change, but which may greatly affect individuals' motivation, choices, effort, and performance. Since supervisory feedback is an important element of educational and HR management strategies - especially with recent technology-driven advances in performance data availability and communication tools - a growing body of literature in psychology, management, and economics analyzes its impact and how it should be given in order to improve performance. Most economic studies focus, however, on the effect of objective performance feedback (e.g. Bandiera et al. 2014, Azmat and Iriberri 2010). Psychological studies devote more attention to the underlying mechanisms and examine how various types of feedback – including subjective feedback - affect motivation and performance (e.g. Deci and Ryan 1985; Locke 1996). In this study, we test whether positive subjective feedback in the form of encouragement and praise, which provides additional information about the environment, has any impact on performance depending on personality traits and gender.

The effect of subjective feedback may be realized through a few channels suggested in previous research. In our experimental game set-up, it may affect individual motivation and *effort*, by (a) affecting the players' confidence or expected performance, and (b) providing utility or disutility. For example, praise may serve as a verbal incentive to perform well, as shown in a recent experiment where workers proved to be more motivated by such communication than a financial reward (Ariely 2016). In the case of praise, only those who perform well receive such a verbal reward. However, to increase the overall performance of a group, it would be important to increase the performance of the underperforming ones as well. Encouragement can be given independently of their performance, and even prior to any effort. It can increase (decrease) a player's utility from the environment, if they prefer friendly communication (or are annoyed by it). It may also increase their confidence and encourage them to choose to make an effort, if their beliefs regarding their expected performance are affected by such unfounded "expressions of faith" from a supervisor. Furthermore, subjective feedback may also impact performance through the players' *marginal productivity*. Some studies have shown that individual productivity may be negatively affected by a stressful environment (Ariely et al. 2009), and positive feedback may improve productivity by setting players at ease. Some studies directly point to gender differences in the effect of praise, for example, females have been shown to have a negative

reaction to praise, due to it being interpreted as "controlling feedback" (Deci et al. 1975).[5] The experimental setup of our study and the data collected allow us to explore the role of these channels, since we observe not only the performance, but also the effort of each player.

Of particular interest to the analysis of gender differences in the response to subjective feedback is the role of self-confidence, which has been shown to differ significantly between the genders and thereby impact decisions and outcomes (Healy & Pate 2011, Wozniak et al 2014). Based on a rudimentary measure of game-specific self-confidence, we test how the effect of subjective feedback varies by confidence level and by gender and confidence level. This allows us to investigate whether any observed gender difference is mediated by the self-confidence channel. If the effect of subjective feedback differs by confidence level, but not by gender within confidence level, then the main implication of our analysis is that the highest benefit from a more targeted communication policy can be achieved based on the targeting of communication by personality type (confidence level) rather than by gender. At the same time, targeting by personality may also improve the relative outcomes of women compared to men, depending on the composition of the particular group (class or workforce).

The results of the experiment point to a significant gender difference in the mean effect of subjective feedback on performance, which is mainly due to gender differences in confidence distributions combined with the differential impact of feedback by confidence level. Females respond positively to encouragement, increasing their performance by around 14%, and negatively to praise, leading to scores that are 9% lower. Males do not show any significant response on average, and are significantly more confident in their game-playing abilities. The interpretation of these results and their implications must be carried out in light of the context. They pertain to a brief game played anonymously in the participant's everyday environment, and therefore reveal real-life decisions, but in a game setting rather than an educational or workplace one. The task (online computer game) may be stereotypically male, which may affect gender differences of behavior. The sample is self-selected, thus it is likely not representative of the entire population in terms of gender or personality traits.

In spite of these caveats, the results do point to some key implications. They show that individuals can react very differently to even small differences in the subjective content of supervisory feedback, depending on their personality traits. The overall effect of a specific feedback on an entire population (class or workforce) is dependent on its composition in terms of gender and personality traits, but the results highlight the importance of paying attention to individual needs rather than relying on standardized communication. Better targeted feedback can improve overall performance and decrease gender differences in

[5] Deci et al. (1975) found that positive feedback increased intrinsic motivation for males, but had the opposite impact for females. The authors explain their findings with the different socialization of males and females, resulting in females focusing on the controlling aspect of positive feedback and responding negatively.

outcomes. The results of our study quantify significant performance losses due to non-optimal feedback, and therefore highlight the economic importance of more individualized treatment. The finding that after controlling for personality, there is no remaining gender difference in the effect of subjective feedback is also remarkable and points to the importance of digging deeper into the causes of gender differences in the experimental literature, as well as more specifically the gender differences of the effect of different types of feedback.

2. THEORETICAL FRAMEWORK AND PREVIOUS EVIDENCE

Our analysis fits into the literature on incentives covering theoretical explanations and empirical evidence from several disciplines. Various forms of incentives have often been analyzed in the economics, management, and psychological literature. Some examples are monetary or verbal rewards, gifts, compliments. We think of subjective feedback as a type of incentive, which may affect individuals' motivation or performance. Additionally, subjective feedback, as defined in our paper, can also affect individuals' output by providing information from the environment, which can be e.g. friendly or rigid yielding directly utility/disutility to participating individuals. The application of various incentives has been a well-analyzed topic. Research in psychology provides a growing number of empirical studies on the relationship between different types of incentives, motivation and performance.

Though incentives are used to improve motivation, the conventionally assumed positive relationship can break down due to several reasons. Cognitive evaluation theory as described by Deci and Ryan (1985) predicts that certain types of rewards might have a detrimental effect on the motivation of individuals, depending on whether the reward is perceived as controlling or informational. Comparing monetary and verbal rewards, they found that monetary rewards decrease the intrinsic motivation of individuals, whereas verbal rewards have no detrimental effect. In a recent experiment, Ariely (2016) found that pizza and compliments proved to be a better motivator of employees than a financial reward. Additionally, the role of individual and demographic characteristics has been suggested as an important factor in how various types of feedback is interpreted (for an overview, see Chang et al. 2012).[6]

[6] For instance, in an early paper Deci et al. 1975 found that positive feedback increased intrinsic motivation for males, but had the opposite impact for females. The authors explain their findings with the different socialization of males and females, resulting in females focusing on the controlling aspect of positive feedback.
In a recent paper, the meta-analysis by Chang et al. (2012) suggests that individuals with higher self-esteem tend to be less „plastic" in their responses to feedback.

The link between incentives and performance depends not only on the relationship between incentives and motivation, but also on how motivation is linked to effort and to performance, which is often described as the impact on individuals' productivity (Ariely et al. 2009). The positive relationship between effort and motivation can be broken e.g. in competitive situations or in the presence of an audience, but personality traits are also important mediators of this relationship (Ariely et al. 2009).

In sum, various psychological theories suggest that incentives, including subjective feedback may affect performance via the following links: (1) incentives have an impact on effort, and (2) effort is related to performance. Both links (1) and (2) may be impacted by individuals' believes in their abilities, or, in general, by their self-efficacy. This is what we call the **belief-updating channel**. Furthermore, individuals may gain utility or disutility from engaging in an activity and investing certain amount of effort. This is what we call as the **utility channel**. Additionally, the effort – performance link may be affected by changes in individuals' productivity, which we call as the **marginal productivity channel**. Link (1) is analyzed e.g. by Deci and Ryan (1985) or Locke (1996), the marginal productivity channel is tested e.g. by Ariely et al. (2009), and the belief-updating channel by Mobius et al. (2011). We add to this literature and separate the role of the two links in determining any gender differences in the effect of subjective feedback. We are, however, unable to differentiate between the belief updating and utility channels. Appendix B provides details of a simple theoretical model used to illustrate the individual's effort decision and the different channels described above, as well as the testable implication regarding gender differences in the effect of subjective feedback when a lower confidence distribution is assumed for women.

Though each of the above theories focuses on different factors of the complex relationship between incentives and performance, they share one common element: individual characteristics play an important role in the link between incentives and effort, and how increased effort is related to performance. In our paper, we take a closer look at the role of individual characteristics in mediating two types of positive verbal feedback: praise and encouragement.[7] In particular, we consider the role of gender and confidence. [8] The praise vs.

[7] Our definition of praise and encouragement has some common elements with the Adlerian notion, therefore, we dedicate a few lines at this point to describe it. The most important distinction between the above two types of positive feedbacks is the presence of evaluation or judgement in case of praise. On the other hand, encouragement aims "to instill courage, perseverance, confidence, inspiration, or hope" (Wong 2015, p. 182) without judging individuals' performance. Another important distinction is that praise evaluates past performance, while encouragement looks into the present or future

[8] It has to be noted though that psychology literature defines confidence and some related terms much more precisely than it is done here. Bandura's theory (Bandura 1977) emphasizes the importance of "self-efficacy", which includes the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p.3). Very simply, self-efficacy is one's belief in his/her abilities to carry out a task. Self-efficacy has an impact not only on performance, but also influences the acceptance of/commitment to goals. Self-efficacy beliefs are determined by past experiences and can be also influenced by "verbal persuasion." Therefore, our research agenda taking a

encouragement debate is also an important topic in parenting and childhood education (Wong 2015), however, related empirical analyses are scarce, especially in the field of economics.[9] Various forms of praise (e.g. monetary vs. verbal rewards) have been frequently the topic of empirical analysis in psychology literature (e.g. Henderlong and Lepper 2002, Deci and Ryan 1999), but encouragement seems to be a less researched theme in most disciplines despite of its substantial practical relevance. Encouragement has been a central component of Adlerian psychology since the 1930's (for an overview, see Wong 2015), and it has becoming lately a promoted element of the leadership management literature (Kouzes and Posner 2003).

The scattered research on the praise vs. encouragement debate, and on the differential effect by personality suggest that encouragement is relatively more important for girls and for minority groups, and the praise vs. encouragement preferences change as kids age (Wong 2015). For instance, Pety et al. (1984) analyzed the preferences of 15 year old and 9-13 year old groups of children regarding praise vs. encouragement, and they found in general that adolescents (15 year olds) have a stronger preference for encouragement than younger kids. Gender differences were also detected, and it has been shown that boys have stronger preference for praise than girls. Kelly (2002) analyzed the praise vs. encouragement preferences of 4th, 6th and 8th grade students as a function their locus of control orientation and gender. The findings suggest that locus of control orientation plays an important role in choosing praise or encouragement. The analysis by gender, similarly to Pety et al. (1984), reveals that boys prefer praise, while girls have a stronger preference for encouragement. One possible explanation proposed by the authors is that boys might feel more comfortable knowing "where they stand", while girls may need more indirect, subtle encouragement. Kelly and Daniels (1997) take a closer look at the acceptance/evaluation of praise and encouragement by school kids, and find significant differences between the praise and encouragement preferences of children: kids uniformly rate encouraging teachers more favorably than praise-giving teachers. Usher and Pajares (2006) found that among sixth-grade students, verbal persuasion is a significant positive predictor of academic self-efficacy for girls and African Americans, but not for boys and non-Latino White Americans.

In sum, research in psychology, especially in the fields of childhood education, provides some empirical evidence on the relative importance of praise vs. encouragement for people with various personal and demographic characteristics. However, the results mostly refer to special groups of individuals (e.g. children in school), include limited number of

closer look at how one's performance may be impacted by positive verbal feedback can also be directly linked to Bandura's model.

[9] The related literature in economics focuses on the role of various firm-level wage-setting schemes to motivate employees (e.g. Lazear 2000), or on the role of objective performance feedback e.g. in the form of exam grade (Bandiera et al. 2014, Azmat and Iriberri 2010, Hannan et al. 2008).

observations, and lack a sophisticated econometric analysis of the results. In our paper, we aim to fill in this gap by targeting the adult population, focusing on performance and effort as outcomes, collecting a larger number of observations and highly detailed data via an online game, and providing quantitative/econometric analysis of the results.

3. METHODOLOGY

3.1. EXPERIMENTAL DESIGN

To obtain the data for testing the theoretical implications empirically, we utilize a simple online computer game.[10] The game is available on a website, where prior to choosing to play, individuals are given a short description of the game, which is simple, but requires concentration and effort. The task is to collect shapes that are displayed in the top left corner of the screen. There are ten different shapes. The shapes move around the screen, and players must find and click on all shapes that match the target shape shown. The game takes 2 minutes, and the goal is to score as many points as possible. No information on other players is given at any point, so the effort and performance of players is not influenced by a public or competitive aspect.

The game is preceded by a simple survey, which asks for basic demographic information (gender, age, location, education level), and data is collected to account for whether the device the game is played on is a touchscreen or not, as this may also affect their final score. Players are also asked to give a nickname that is used to identify games played by the same person. The survey also includes questions related to the individual's own experience with games (plays often, sometimes, never), and to their task-related confidence, asking how good the individuals consider themselves to be at computer games (excellent, pretty good, ok, pretty bad, very bad). The survey is designed to be as quick and easy to fill out as possible, and to not raise any suspicions regarding the purpose of the experiment. We therefore use this relatively rudimentary measure of confidence in our analysis. It may be argued that this measure does not truly capture individual confidence, however, we believe that it is relevant to our analysis in that it closely reflects the individual's beliefs regarding how well they will play the game, just prior to playing it, and thereby have an effect on their motivation and effort. The results indeed confirm that it plays a significant role in determining the effect of the subjective feedback, and that it varies highly between the genders, in line with previous empirical evidence.

[10] The game is available on a website: https://experimental-games.herokuapp.com/#bw.

When players click to start playing the game, they are randomly selected to receive one of the following treatment options. The four options represent praise, encouragement, and a corresponding control group for each:

- **Control (Praise)**: Performance feedback (the player is given information on the number of targets completed, after given numbers of completed targets)
- **Control (Encouragement)**: Performance feedback (the player is given information on their score, at given times)
- **Praise**: Performance feedback + Praise (the player is given information on the number of targets completed, together with praise in the form of graphics and text, after given numbers of completed targets)
- **Encouragement**: Performance feedback + Encouragement (the player is given information on their score, together with encouragement in the form of graphics and text, at given times)

Positive subjective feedback is given in the form of text (such as "Good job!" or "You can do it!") and graphics (simple, culturally neutral emoticons and pictures). The phrases in the praise treatment contain positive valuations of past performance, while those in the encouragement treatment refer to expectations of future performance and expressions of support. Appendix A gives the specific details (timing, text, graphics) for each treatment type. The texts and pictures were chosen to be as international as possible in order to clearly convey the same meaning. However, they may still evoke very different reactions from different individuals, which is exactly what we expect, and why we think that the treatment effect may differ by personality and gender. For example, males may be more likely to find the emoticons childish and annoying, while females may be more likely to perceive them as friendly and pleasant, and these perceptions may differ by culture.[11] The estimated effects are meant to include such differences in preferences, as well as differences in how the feedback may affect players' confidence. However, it is important to note that all of the results pertain to only these specific treatments. The feedback is communicated via pop-up textboxes, which have to be clicked for the game to continue. After each treatment given during the game, the shapes that are displayed on the screen are "shaken up." These features were implemented to ensure that the subjective feedback is noticeable, but at the same time, it does not seem intrusive or out of place.

[11] Comments received from game testers during follow-up interviews were very much in line with this: several males mentioned that the emoticons and phrases were annoying, while several female testers mentioned how friendly the design and communication was.

Two control groups are included in order to account for the different timing of the feedback types: praise is given after a certain number of successfully completed targets, while encouragement is given independently of performance, at given points in time (see appendix A). This is a key difference between the two treatments, which means that the effect of praise may be more dependent on performance for two reasons: the frequency of treatment will depend on the player's performance, and the perception of it may depend on it as well. Players who feel that they are doing well will regard the praise as genuine, while those who feel they are doing badly may see it as mocking or simply false.[12] To ensure that the design of the game is the same for control and treatment groups, the shapes displayed on the screen are also "shaken up" after receiving information on the number of completed targets for control group players. It is important to note that players receive the same feedback if they play again within each gaming session, but not necessarily between sessions. We therefore limit our analysis to the results of the first session of each player, which means that we are able to analyze the effect of treatment over a length of one game to a maximum of a few games.

The output data is highly detailed, recording every event that takes place during the game (clicks, treatments) to the thousandth of a second, along with the data from the pre-game survey. This allows us to analyze not only the end outcomes of each game, but the evolution of effort and performance over time as well. The empirical analysis therefore relies on the comparison of group-level means of the main outcome variables (end score and total number of clicks[13]) and corresponding t-tests, and the graphical representation of the evolution of outcomes over time by groups.

It is important to note how the experimental design affects the interpretation and external relevance of the results. The experiment is based on a game as the task, played at home or in the setting where game-playing usually takes place, and is completely anonymous and with no public or social aspect. In some sense, it therefore provides evidence of real life behavior in a natural setting, but not an educational or labor market one. The results may also be task-specific: gender differences are generally smaller in tasks perceived as stereotypically more female (Niederle 2016), and game-playing is more likely more typically male, as confirmed by the pre-survey questions pertaining to game playing frequency. As mentioned earlier, any effects measured are also specific to the particular treatment specification. Feedback is given through clearly pre-programmed communication rather than a supervisor, using specific emoticons and simple, commonly known phrases. These impact

[12] This was also suggested by several testers in the follow-up interviews.
[13] We plan to further extend our outcome measures to include the number of times players play again within the first session, and whether they return for another session, which will capture slightly longer-run effects.

both the perception of the feedback and its effect, as individual reactions would likely differ in the case of more serious, personal praise received a respected supervisor.

Another limitation to the external relevance of the results may stem from the sampling method used. In our main sample, the game link is shared using social media outlets, and participation is voluntary, so the sample includes those who were willing to play a game to support research based on a social media ad, and who are probably not representative of the full population in either gender or traits. Arechar et al. (2017) discuss the benefits and problems of online experimenting,[14] and conclude that data collected from online experiments is reliable and can provide the basis for valuable contributions to the empirical evidence. They also highlight the potential selection bias due to participant dropouts as the most important issue. Of course, the importance of such selection is also dependent on the sampling method. In order to assess the size of this issue in our case, we also collect data in laboratory settings with university students as the participants, and compare the results in order to test the sensitivity of our estimates to selection in participation. Of course, the more controlled classroom sample suffers from its own drawbacks: it is limited in terms of the age, region, and education level of the participants.

3.2. EMPIRICAL TESTS

We first test whether we can observe any significant impact of subjective feedback on effort or performance for the full sample of individuals. In line with reviewed theories, performance may be positively affected by either increased effort or marginal productivity. We compare the outcomes of the control and treatment groups separately for praise and encouragement, first looking at performance and then at effort. The mean population effect depends on the effect on different sub-groups (e.g. gender) and the ratio of those sub-groups within the population (sample). A finding of a significant positive (negative) impact would suggest that giving positive subjective feedback to everyone would have a positive (negative) overall impact on performance. However, even if no significant impact is found for the full sample, it is possible that more targeted feedback could improve overall outcomes by improving the performance of certain sub-groups.

To assess whether feedback targeted by gender could improve overall performance, we move on to our main question, that is, we test whether the effect of positive subjective feedback differs by gender. Treatment will increase effort more for women if they derive a

[14] The authors discuss the available evidence and carry out a well-known experiment using both a laboratory setting and the online labor market platform Amazon Mechanical Turk. They find that basic behavioral patterns are replicable online, and conclude that online experiments can provide data of adequate quality and are a valuable complement to laboratory studies.

higher utility from subjective feedback, or they incorporate it with a higher weight in their belief updating, as suggested in section 2. Performance may also be affected through the third channel of marginal productivity: if the marginal productivity of women increases more due to treatment than that of men, we can expect to find a positive effect of treatment. To test for gender differences, we first compare the difference in the performance of the control and treatment groups by gender, and test for significant treatment effects by gender, as well as for significant differences in the treatment effect between women and men. We again assess whether any significant treatment effects arise through an impact on effort, or if they are due to changes in marginal productivity, by comparing the impact on performance to that on effort.

The hypothesis that females may be more affected by positive subjective feedback is motivated by previous research on gender differences in personality traits (Feingold 1994, McCarty 1986, Mobius et al. 2011), which finds that women are on average less self-confident than men. Since feedback may serve as a tool to counteract unfavorable individual or environmental conditions (e.g. lower self-confidence or high pressure), we expect differential effects by gender. Assuming that females are on overall less confident than men, receiving treatment will increase effort more for women if positive subjective feedback increases utility more for those with lower self-confidence, or if the self-evaluation of those with lower confidence is affected more by the feedback. Performance may also be affected through a gender difference in the effect of feedback on marginal productivity, i.e. if positive feedback increases the marginal productivity of less confident individuals more than that of more confident ones, for example, by making them feel more comfortable in their environment. To assess the role of self-confidence, we first compare treatment effects in the overall sample by confidence level, and then examine them by gender and confidence level. We also examine whether women have lower confidence on average in our sample.

Finally, we also assess how the effect of praise differs from that of encouragement. To our knowledge, this question has not been analyzed in the literature so far. The effect may differ for a few reasons. First, it is important to note how their information content differs. While both are likely to be perceived to contain similar information about the friendliness of the environment relevant to the direct (dis)utility, praise is more likely to be perceived as information on the player's own performance that is relevant to belief updating. Praise contains information about past performance and can only be given based on performance,[15] while encouragement refers to future performance, and can be given independently of performance. These types of information may be weighted differently in the updating of beliefs, and in consequence may have different effect on individual effort and performance.

[15] This is not necessarily true, as one can also praise effort instead of performance. However, we are only focusing on praise given for good performance here.

Second, the magnitude of treatment also differs: since praise is by definition dependent on performance, high performers receive more treatment, while encouragement can be (and is, in our case) given independently of performance. The differences by performance level mean that praise can actually exacerbate existing inequalities, by aiding those who are already performing well, and having a smaller effect on those who are not due to less treatment, or even a negative effect if it is perceived as mocking. On the other hand, encouragement may be an important tool for aiding weaker performers and decreasing inequalities, as it can be given independently of performance.[16]

4. RESULTS

Table 1 summarizes the available characteristics of our on-line sample. The sample consists of 343 individuals, out of which 199 are men and 142 are women, and 602 games played. The sample is comprised of mostly highly educated individuals of the 20-40 age group, which points at the potential problem of non-representativeness driven by the sampling methods. Furthermore, despite the fact that we collect data on individuals coming from various countries, the sample is dominated by Hungary (58%) and Poland (20%).

Table 1

Basic statistics of the sample

	Females			
	Control (Enc.)	Control (Praise)	Encouragement	Praise
N	32	36	44	32
Age	30.1	30.2	29.5	31.2
Education	1.9	1.8	1.8	1.9
	Males			
N	47	52	50	50
Age	29.5	32.5	29.6	29.5
Education	1.8	1.8	1.8	1.8

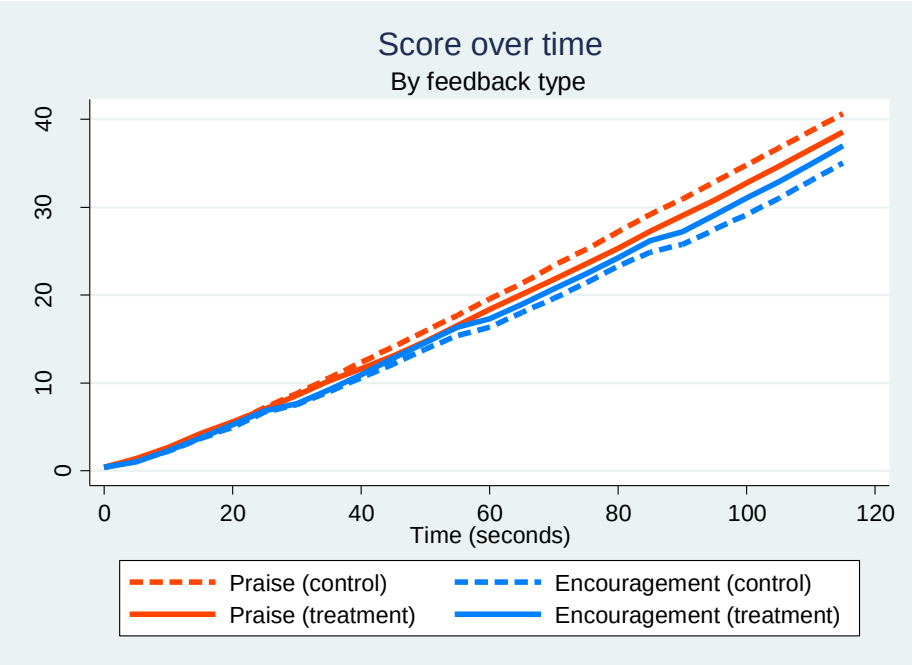
We begin the discussion of the empirical results with the estimated effects of the two feedback types on the full sample of individuals (pooled men and women). Figure 1 depicts the evolution of mean scores by treatment type over time, during the full length of the two minute game. It shows no evidence of significant differences due to treatment, as confirmed by t-tests performed on mean differences in outcomes by treatment type. As expected, due to the differences in timing, the scores of those receiving Control (Praise) are slightly higher

[16] Encouragement could also be targeted towards weak performers specifically, as long as it is not perceived to be so, which would lead to a negative effect similar to that of praise.

than those receiving Control (Encouragement), since the former involves text boxes appearing and shapes being "shaken up" during target shape changes, while the latter involves random interruptions that are more disruptive to game playing. The pooled results suggest that giving subjective feedback to everyone would not change the mean overall outcome in any significant way.

Figure 1

Evolution of scores by feedback type, full sample



We next turn our attention to gender differences in the impact of feedback. We analyze the evolution of the treatment effects, defined as a difference in the mean score of those receiving a treatment (either encouragement or praise) and respective control group. Figure 2.a plots the treatment effects by feedback type and by gender over time. The data suggest significant differences in the effect of subjective feedback - both by gender and by feedback type (encouragement vs. praise). We observe, that regardless the feedback type females are affected by receiving feedback more strongly than men, for whom it does not impact performance. .Table 2 additionally gives the mean end scores and their differences after receiving a treatment (either encouragement or praise). T-tests confirm that these treatment effects – both for encouragement and for praise - are significant in the case of women, but not in the case of men. More specifically, they show that females are affected positively by receiving encouragement, achieving scores that are 4.5 points (14%) higher on average. The estimated effect of praise is negative (-3.5 points or -9%), but slightly below significance.

Figure 2.b. depicts the evolution of the treatment effect calculated for the number of clicks (a difference in the mean number of clicks of those receiving a treatment (either encouragement or praise) and respective control group), which capture individuals effort. The relevant means and T-tests are indicated in Table 2. The results suggest that the positive effect of encouragement among females appears to be realized through higher effort. Praise in turn, has a significant and large negative impact on the effort among females, which is not entirely realized in their performance. Interestingly, while men's effort is not affected by encouragement, similarly to their performance, the results for praise show a different picture. Their effort increases significantly, by 5.3 clicks, while their performance remains unaffected, suggesting that their marginal productivity (i.e. accuracy in clicking the correct shapes) decreases due to treatment at the same time.

Figure 2.a

Treatment effect on performance over time, by gender

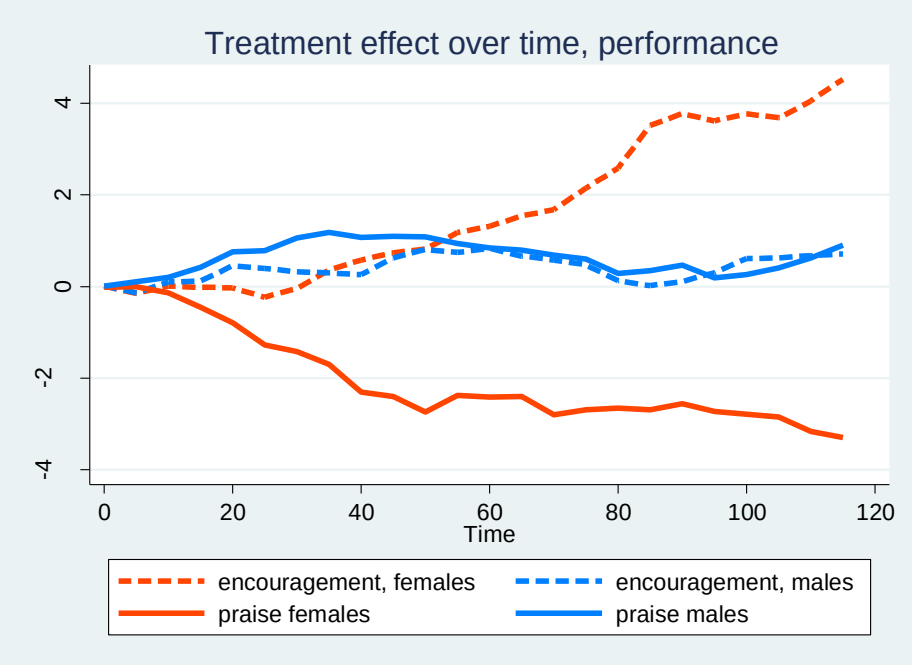


Figure 2.b

Treatment effect on effort over time by gender

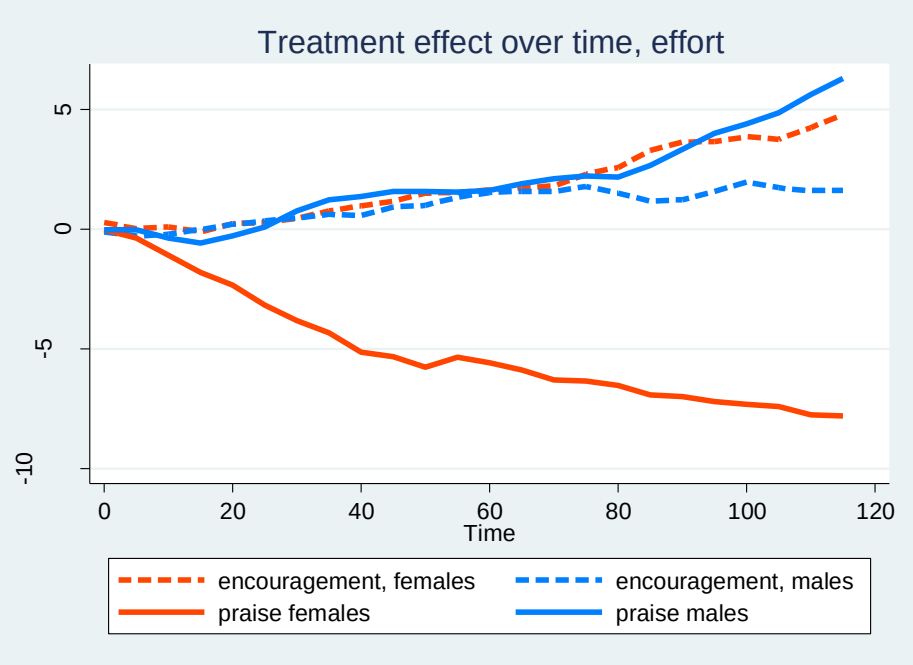


Table 2

Mean treatment effects on performance and effort by gender

		Females		Males	
		Performance	Effort	Performance	Effort
Encour-agement	Control	31.5	43	32.8	45.9
	Treatment	36	47.7	32.9	46.7
	Difference	4.5	4.7	0.1	0.8
	P-value	0.07	0.09	0.47	0.41
Praise	Control	38.6	52.6	37.4	46.8
	Treatment	35.1	44.7	37.8	52.1
	Difference	-3.5	-7.9	0.4	5.3
	P-value	0.14	0	0.45	0.07

We next turn our attention to the role of confidence as the mediator in these effects. Table 3 summarizes the distribution of observations by gender and confidence level overall and by feedback type. The self-reported task-specific confidence measure shows significant differences by gender: only 6 percent of females report themselves as being good at playing

games, while 46 percent report being bad at it. For males, the distribution is opposite: only 11% report being bad at games, and 39 percent consider themselves to be good.

Table 3

Distribution of confidence by gender and treatment type

	Females					
Confidence	Total	% total	Control (encouragement)	Control (praise)	Encouragement	Praise
Bad	127	46%	20	40	43	24
OK	131	48%	27	39	40	25
Good	16	6%	6	2	3	5
All	274		53	81	86	54
	Males					
Confidence	Total	% total	Control (encouragement)	Control (praise)	Encouragement	Praise
Bad	36	11%	2	10	9	15
OK	165	50%	44	61	30	30
Good	127	39%	32	18	34	43
All	328		78	89	73	88

Table 4 summarizes the means and treatment effects by feedback type, gender and confidence level. For the pooled sample of women and men, the results suggest that encouragement has a significant positive effect for those at the middle confidence level, and a negative effect on those with high confidence in their gaming abilities. Turning to the results by gender and confidence level, we can see which subgroups drive the results by confidence level, and whether gender differences exist within confidence categories. The positive effect of encouragement within the middle confidence level is common among both women and men. On the other hand, the negative effect on high confidence players appears to be driven by the effect on men, however, as noted earlier, the number of observations in some of the gender-confidence cells is very low, in particular, we observe very few highly confident females, so the results should be interpreted with this in mind. Therefore, the evidence is not indicative of gender differences in the effect of encouragement within confidence levels. The effect of praise clearly depends negatively on confidence level, likely due to the issue of perception of the feedback discussed earlier: it has a significant negative impact on those with low confidence, no impact on the middle category, and a positive significant impact on highly confident players. Again, a comparison of the effects by gender does not reveal any differences within confidence level. The results for effort do not point to any of the effects on

performance being realized through changes in marginal productivity, as they generally mirror the trends we just described.

Table 4

Treatment effect on performance by confidence level

Performance (end score)						
	Treatment effect: encouragement			**Treatment effect: praise**		
Confidence	**Overall**	**Women**	**Men**	**Overall**	**Women**	**Men**
Bad	0.9	0.7	2.6	-13.9**	-12.6**	-12.1**
OK	8.6**	8.9**	7.9**	2.1	5.5	-0.3
Good	-7.9**	2.7	-9.1**	10.3**	10.2	10.3**
Effort (number of clicks)						
	Treatment effect: encouragement			**Treatment effect: praise**		
Confidence	**Overall**	**Women**	**Men**	**Overall**	**Women**	**Men**
Bad	0.9	-0.52	10.9	-17.6**	-19.0**	-8.8**
OK	11.1**	14.7**	8.3**	1.0	2.5	-0.2
Good	-1.4**	-14.7	-8.8**	18.6**	9.7	19.6**

Though our results are not directly comparable to empirical evidence from previous papers, they suggest similar conclusions regarding the interplay of individual characteristics and feedback. Encouragement, as opposed to praise, has found to be a better motivator of minority groups and females in earlier research in education and psychology (Wong 2015). Deci et al (1975) explain their results with the different socialization of males and females, while the findings of Zeldin and Pajares (2000) suggest that women value the opinion of other people more highly. This is exactly what our results also suggest: women tend to respond positively to encouragement, while males' performance is not affected by it.

5. CONCLUSION

In this paper, we study the effect of positive subjective feedback, in the form of encouragement and praise, on the effort and performance of individuals, and specifically, the gender differences in this effect. We carry out our analysis using data gathered from an online game, in which players are randomly chosen to be given either no subjective feedback (control), or positive subjective feedback (treatment). We analyze the performance (score)

and effort (number of clicks) of players by feedback type, gender, and self-reported confidence level, using data collected from both classroom experiments and online users. The results point to a significant gender difference that is related to self-confidence, which has important implications for gender differences in school and labor market outcomes and efficiency.

Evidence of such differences suggest that one reason for existing gender inequalities – for example, the gender wage gap or the gap in employment in higher level positions – may be that current environments utilize communication that is, on average, better suited to the needs of males. Better targeted feedback, such as encouragement given to under-confident females, could increase the performance of such individuals, and consequently, overall performance. It is important to note that these findings do not mean that all forms of encouragement and praise would have a similar effect, their external relevance is limited for several reasons. The important result is that even within such a brief task and based on very small changes to the communication received during that task, we see significant differences in the response to the same feedback among gender and confidence groups. This suggests that the effects may be significantly higher in a real life supervisory relationship that involves much more communication over longer time periods. The results therefore suggest that significant economic losses may occur due to non-optimized, untargeted supervisory communication, for example, due to lower confidence individuals not receiving sufficient encouragement. Much more future research is needed to develop truly relevant practical recommendations on this topic.

REFERENCES

Altonji, J., Blank, R. (1999). Race and gender in the labor market. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3C. Elsevier, Amsterdam, pp. 3143 – 3259.

Ariely, D., U. Gneezy, G. Loewenstein and N. Mazar (2009), Large Stakes and Big Mistakes, Review of Economic Studies 76, 451-469.

Azmat, G., C. Calsamiglia and N. Iriberri (2015), Gender Differences in Response to Big Stakes, CEP Discussion Paper No. 1314.

Azmat G. and B. Petrongolo (2014), Gender and the labor market: What have we learned from field and lab experiments?, Labour Economics 30, 32-40.Bandiera, O., V. Larcinese and I. Rasul (2014), Blissful Ignorance? A Natural Experiment on the Effect of Feedback on Students' Performance, Working Paper Nr. 511, IGIER – Università Bocconi.

Bandura, A. (1996). Self-efficacy: The exercise of control. New York: Freeman.

Baumeister, R. F. (1984), Choking Under Pressure: Self-Consciousness and Paradoxical Effects of Incentives on Skillful Performance, Journal of Personality and Social Psychology, 46 (3): 610-20.

Bertrand, M. (2011). New Perspectives on Gender. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 4B. Elsevier, Amsterdam, pp. 1545-1592. Chang, Ch. D. Lance Ferris, Russell E. Johnson, Christopher C. Rosen, James A. Tan (2012), Core Self-Evaluations: A Review and Evaluation of the Literature, Journal of Management38 (1), 81-128.

Croson, R. and U. Gneezy (2009), Gender Differences in Preferences, *Journal of Economic Literature* 47 (2), 448–474.

Deci, E.L. (1971), Effects of externally mediated rewards on intrinsic motivation, Journal of Personality and Social Psychology 18, 105–115.

Deci, E.L., Cascio, W.F., and Krusell, J. (1975), Cognitive evaluation theory and some comments on the Calder and Staw critique, Journal of Personality and Social Psychology, 31, 81–85.

Deci, E. L., Koestner, R., and Ryan, R. M. (1999), A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation, Psychological Bulletin, 125, 627-668.

Deci, E. L., and Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum.

Deci, E. L., and Ryan, R. M. (2000), The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior, Psychological Inquiry11 (4), 227–268.

Duckworth, A. L., Eichstaedt, J. C. and Ungar, L. H. (2015), The Mechanics of Human Achievement. Soc. Personal. Psychol. Compass 9, 359–369.

Eckel, C. C. and P. J. Grossman (2008) Differences in the Economic Decisions of Men and Women: Experimental Evidence. In: Plott, Ch. R., . Smith, V. L. (Eds.), Handbook of Experimental Economics Results, Vol. 1, 509-519.

Feingold, A. (1994), Gender differences in personality: a meta-analysis, Psychol Bull., 116(3), 429-56.

Goldin, C., L. Katz, and I. Kuziemko (2006), The Homecoming of American College Women: The Reversal of the Gender Gap in College, Journal of Economic Perspectives 20, 133-156.

Hannan, R.L., R. Krishnan and A. H. Newman (2008), The Effects of Disseminating Relative Performance Feedback in Tournament versus Individual Performance Compensation Plans, SSRN Electronic Journal 83(4):893-913.

Henderlong, J., M. R. Lepper (2002), The Effects of Praise on Children's Intrinsic Motivation: A Review and Synthesis, Psychological Bulletin 128(5), 774–795.

Iriberri, N. and G. Azmat (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.

Katz, I., A. Assor, Y. Kanat-Maymon and Y. Bereby-Meyer (2006), Interest as a motivational resource: Feedback and gender matter, but interest makes the difference, Social Psychology of Education 9, 27–42.

Kelly, F. D., and Daniels, J. G. (1997), The effects of praise versus encouragement on children's perceptions of teachers. The Journal of Individual Psychology 53, 1-11.

Kelly, F. D. (2002), The effects of locus of control, gender, and grade upon children's preference for praise or encouragement. The Journal of Individual Psychology, 58, 197-207.

Kouzes, J. M. and B. Z. Posner (2003). Encouraging the Heart: A Leader's Guide to Rewarding and Recognizing Others. Jossey-Bass, A Wiley Imprint.

Lazear, E. P. (2000), Performance pay and productivity, American Economic Review 90, 199-236.

Locke, E. A. (1996), Motivation through conscious goal setting, Applied & Preventive Psychology 5, 117-124.

McCarty, P. (1986), Effects of Feedback on the Self-Confidence of Men and Women, The Academy of Management Journal 29 (4), 840-847.

Mobius, M. M., M. Niederle, P. Niehaus, T. S. Rosenblat (2011), Managing Self-Confidence: Theory and Experimental Evidence, NBER Working Paper 17014.

Niederle, Muriel, (2016). Gender, In: Kagel, J. and oth, A.E. *Handbook of Experimental Economics*, second edition, Princeton University Press, 481-553.

Pety, J., Kelly, F. D., and Kafafy, A. E. (1984), The praise–encouragement preference scale for children. Individual Psychology: Journal of Adlerian Theory, Research & Practice 40, 92-101.

Usher, E. L., and Pajares, F. (2006), Sources of academic and self-regulatory efficacy beliefs of entering middle school students. Contemporary Educational Psychology 31, 125-141.

Wong, J. Y. (2015), The Psychology of Encouragement, The Counseling Psychologist 43(2), 178 – 216.

Zeldin, A. L. andPajares, F. (2000), Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. American Educational Research Journal 37, 215-246.

APPENDIX A

Table A1

Details of the treatment specifications

Treatment type								
Control (Praise)		**Control (Encouragment)**		**Praise**			**Encouragement**		
Trigger	**text**	**Trigger**	**text**	**Trigger**	**Picture**	**text**	**Trigger**	**Picture**	**text**
start screen	Are you ready? "Click"	start screen	Are you ready? "Click"	start screen	x	Are you ready? "Click"	start screen		Are you ready? "Good luck!" "Click"
after 2nd, 5th, 10th, 15th, 20th... shape change	"X targets conpleted"	30/60/90 seconds	Score: X	after 2nd, 5th, 10th, 15th, 20th... shape change		"X targets completed" + 3 texts alternate: "Good job!" or "Well done!" or "You're great!"	30		Score: X + "You can do it!"
END	Score: XX Play again!	END	Score: XX Play again!	END		Congratulations! Score: XX Play again!	60		Score: X + "Keep it up!"
							90		Score: X + "Almost there!"
							END		Score: XX + Play again!

APPENDIX B: THEORETICAL MODEL

We propose a simple model based on goal-setting theory (Locke 1996), in order to illustrate the links between personality, gender, and the response to supervisory feedback, and develop clear testable implications regarding the effect on effort and performance. We focus on the information content of supervisory feedback, distinguishing between subjective and objective information. Let I_o refer to objective information about the individual's performance, i.e. *performance feedback*. For example, in a game, players may be given their score or ranking. Respectively, let I_S refer to subjective information about an individual's performance that is received from the environment (supervisor).

We further distinguish between the two types of subjective information studied in our experiment. I_{SP} contains subjective information about the player's past performance, i.e. *praise*, and is given after some type of success occurs. Examples include phrases such as "Good job" or "Well done." I_{SE} refers to subjective information about the player's expected future performance, i.e. *encouragement*, and may be given independently of performance, or even prior to the game. Examples include phrases such as "You can do it" or "Keep it up." The three types of information, i.e. I_o, I_{SP}, I_{SE} , received during the game may affect the individual's expectations of the environment, as well expectations of their own performance, and these effects may differ among individuals. We focus on whether they differ by gender and confidence level.

We first present a simplified model based on a single type of subjective feedback, and differentiate among the two types (I_{SP} and I_{SE}) later on. In our model, individuals choose effort (e) to maximize their utility:

$$\max_e E(U) = \{\Psi + \Gamma(S)\} * e \quad (1)$$

where effort can be discrete, meaning whether an individual plays or not, or continuous (e = {0; 1}), expressing how hard an individual tries. Expected utility from playing the game is composed of two parts: Ψ, a success-independent part, and $\Gamma(S)$, a success-dependent part. S, or success, is defined as reaching the goal previously set by the individual, for example, „Do the best you can," or "Beat my previous high score." S takes the value of 1 if the goal is achieved, and 0 otherwise. Goal setting itself varies among individuals, and may be correlated with gender or confidence. Note that in the current experimental setup, there is no information given to players about the performance of others, so we are eliminating the possibility of competitive goal setting.

The success independent utility (or disutility) derived from playing by individual i, Ψ_i, is of the following form:

$$\Psi_i = E_i(X_i) * R_{X,i} + I_S * R_{IS,i} \quad (2)$$

where X_i includes any factors that increase or decrease the individual's expected utility but are independent of success. Such factors include the time and energy spent on playing the game (decreasing utility), as well as learning from the game or enjoyment gained from playing the game (increasing utility). $R_{X,i}$ reflects individual preferences and quantifies the returns the individual draws from factors in X_i, which may be positive or negative. Subjective feedback, I_S, may serve as a direct source of utility (or disutility). For example, a supportive environment may make the game more enjoyable for some individuals, or may be seen as annoying by others and decrease utility. $R_{IS,i}$ represents the return to subjective feedback, which may be positive or negative, and varies at the individual level.

The success-dependent part of the individual's expected utility, $\Gamma_i(S)$, is of the following form:

$$\Gamma_i(S) = E_i(S) * E(R) - \big(1 - E_i(S)\big) * E(L) \quad (3)$$

where $E_i(S)$ represents the individual's expected probability of success. $E(R)$ contains the expected rewards of success, such as: any rewards, a sense of achievement, public pride, or increased self-confidence. On the other hand, $E(L)$ captures the expected losses stemming from failure in achieving the goal, such as: punishment, sense of failure, public shame, or decreased self-confidence. Note that in our current setup, there is no public, and no reward or punishment is given in the game, so these returns are not included. An individual's expectation of success, $E_i(S)$, depends on the individual's baseline self-confidence (C_0), any information received by the individual about the performance received previously (I_o, I_S), and the individual's weighting of the information ($w_{IO,i}$ and $w_{IS,i}$), which varies among individuals:

$$E_i(S) = C_{0,i} + I_o * w_{IO,i} + I_S * w_{IS,i} \quad (4)$$

Substituting these back into the expected utility equation (1) gives:

$$E_i(U) = \{\Psi + \Gamma(S)\} * e = \begin{Bmatrix} E(X_i) * R_{X_i} + I_S * R_{IS} + \\ +\big(C_0 + I_{o,t-1} * w_{IO} + I_{E,t-1} * w_{IE}\big) * E(R) - \\ -\big(1 - \{C_0 + I_{o,t-1} * w_{IO} + I_{E,t-1} * w_{IE}\}\big) * E(L) \end{Bmatrix} * e \quad (5)$$

Individuals will choose to participate in the task, i.e. undertake an effort and choose $e^* = 1$ if $\varphi = \frac{\partial E(U)}{\partial e} \geq 0$. The effect of subjective feedback on optimal effort is then given by:

$$\partial\varphi / \partial I_S = R_{IS,i} + w_{IS,i} * \big(E(R) + E(L)\big) \quad (6)$$

The choice of effort can therefore be influenced by subjective feedback through two channels: *direct utility returns*, depending on individual environmental preferences ($R_{IE,i}$); and *belief updating*, depending on the individual-specific weighting of subjective information received in the expectations of success ($w_{IS,i}$).

The performance (O) of each individual is given by:

$$O_i = MP_i * e_i \tag{7}$$

where MP_i represents the marginal productivity of individual i. Besides that subjective feedback may have an effect on individual performance through the individual's choice of effort, it is also possible that players' marginal productivity may be affected. For example, some individuals may be prone to make more mistakes when they feel stressed or nervous, but they improve their accuracy in a task when they feel more comfortable due to supervisory communication that sets them at ease. Therefore, there is a third pathway through which subjective feedback may affect performance: the *marginal productivity* channel.

In our model, R_{IE} and w_{IE} may differ along personality types (p):

$$\partial\varphi/\partial I_E = R_{IE}(p) + w_{IE}(p) * \big(E(R) + E(L)\big) \tag{8}$$

We focus on self-confidence and define p as the following:

$$p = \begin{cases} 0 \; for \; low \; type: less \; confident \\ 1 \; for \; high \; type: more \; confident \end{cases} \tag{9}$$

Based on the previous literature, this gives us:

$$\frac{1}{n}\sum_{i=1|i\,female}^{n} p_i < \frac{1}{k}\sum_{j=1|j\,male}^{k} p_j \tag{10}$$

Receiving treatment I_E will increase effort more for women if either $\frac{\partial R_{IE}(p)}{\partial p} \leq 0$ and $\frac{\partial w_{IE}(p)}{\partial p} \leq 0$ holds, with at least one of them being nonnegative, i.e. if positive subjective feedback increases utility more for those with lower self-confidence, or if the self-evaluation of those with lower confidence is affected more by the feedback. Performance may also be affected through a gender difference in the effect of feedback on marginal productivity, i.e. if positive feedback increases the marginal productivity of less confident individuals more than that of more confident ones, by making them feel more comfortable in their environment.